
**U.K. HOUSE OF COMMONS ELECTIONS:
CANDIDATES, CONSTITUENCIES,
AND CAMPAIGN SPENDING
1885 - 2019**

October 27, 2020

Alexander Fournaies
Harris School at University of Chicago

Contents

1	Introduction	3
2	Description of Variables	5
3	Citation	17
4	Appendix	19

Acknowledgements

I want to thank David Butler, Gary Cox, Andy Eggers, Iain McLean, Michael Pinto-Duschinsky, and Arthur Spirling.

Chapter 1

Introduction

The dataset `commons_candidate_campaign_spending` contains information on UK House of Commons candidates, constituencies, and campaign spending from 1885 to 2019. This codebook describes the dataset.

The dataset comprises the complete universe of parliamentary candidates from 1885 to 2019. Each row in the dataset pertains to a parliamentary candidate in a particular election, and the variables contain constituency- and candidate-level information.

A long list of prominent scholars have collected historical data on UK House of Commons elections and made it publicly available. This project builds on their work, either directly or indirectly, but extends it in at least three important ways.

First, the dataset contains detailed, candidate-level information on the level and composition of campaign expenditure for more than 99.7% of all general-election candidates. In addition to information on the level of permitted spending, the dataset documents how much money each candidate allocated to advertisement, campaign managers, staff, public meetings, facilities, miscellaneous matters, and personal expenses.

Second, the dataset includes candidate identifiers that track individual candidates over time, even if they change name or constituency. These identifiers are constructed using various sources of biographical information.

Third, the dataset includes constituency identifiers that track constituencies from one boundary change to the next. These identifiers are constructed on the

basis of historical maps and detailed descriptions of constituency boundaries.

Chapter 2

Description of Variables

Each row in the dataset is uniquely identified by a parliamentary candidate in a given election. The dataset contains 62 variables. Each variable is briefly described below.

<code>obs_id</code>	Numeric variable. Discrete variable that uniquely identifies the observations in the dataset. The variable has 78,548 non-missing observations and 78,548 distinct values.
<code>nation</code>	Numeric variable. Discrete variable that reports whether the parliamentary constituency is located in England (<code>==1</code>), Ireland (<code>==2</code>), Northern Ireland (<code>==3</code>), or Scotland (<code>==4</code>), or Wales (<code>==5</code>). The variable has 78,548 non-missing observations and 5 distinct values.
<code>region</code>	Numeric variable. Discrete variable that reports whether the parliamentary constituency is located in East England (<code>==1</code>), East Midlands (<code>==2</code>), Greater London (<code>==3</code>), Ireland (<code>==4</code>), North East England (<code>==5</code>), North West England (<code>==6</code>), Northern Ireland (<code>==7</code>), Scotland (<code>==8</code>), South East England (<code>==9</code>), South West England (<code>==10</code>), Wales (<code>==11</code>), West Midlands (<code>==12</code>), or Yorkshire and the Humber (<code>==13</code>). The variable has 78,548 non-missing

observations and 13 distinct values.

<code>const</code>	String variable. String variable that reports the name of a parliamentary constituency, the period during which it was in use, and its classification (borough, county, or university). The variable has 78,548 non-missing observations and 4,048 distinct values.
<code>const_alternative</code>	String variable. String variable that reports the name of a parliamentary constituency, the period during which it was in use, and its classification. The variable has 78,548 non-missing observations and 4,506 distinct values.
<code>const_category</code>	String variable. String variable that reports whether the constituency is classified as a borough, county, or university constituency. The variable has 78,548 non-missing observations and 3 distinct values.
<code>year</code>	Numeric variable. Discrete variable reporting the year of the election. The variable has 78,548 non-missing observations and 133 distinct values.
<code>date</code>	Numeric variable. Variable reporting the date of the election in question. The variable has 78,548 non-missing observations and 1,642 distinct values.
<code>by_election</code>	Numeric variable. A dummy variable indicating whether the election is a by-election (<code>==1</code>) or a general election (<code>==0</code>) The variable has 78,548 non-missing observations and 2 distinct values.
<code>seats_up</code>	Numeric variable. Discrete variable that reports the number of seats that the constituency returns to the House of Commons in the election in question. The variable has 78,548 non-missing observations and 3 distinct values.

<code>seats_total</code>	Numeric variable. Discrete variable that reports the number of MPs who represent the constituency in question. This may differ from the variable <code>seats</code> if not seats are up for reelection in a by election. The variable has 78,548 non-missing observations and 3 distinct values.
<code>electors</code>	Numeric variable. A discrete variable reporting the number of electors in a constituency. The variable has 75,772 non-missing observations and 18,848 distinct values.
<code>turnout</code>	Numeric variable. A continuous variable reporting the turnout in the election. Turnout is defined as $\frac{\text{totalvotes}}{\frac{\text{seats}}{\text{electors}}}$. The variable has 74,261 non-missing observations and 22,438 distinct values.
<code>candidates</code>	Numeric variable. Discrete variable that reports the total number of candidates who run for a parliamentary seat in a given election. The variable has 78,548 non-missing observations and 20 distinct values.
<code>polling_districts</code>	Numeric variable. A discrete variable reporting the number of polling district in a constituency. The variable has 48,971 non-missing observations and 176 distinct values.
<code>polling_stations</code>	Numeric variable. A discrete variable reporting the number of polling stations in a constituency. The variable has 48,835 non-missing observations and 199 distinct values.
<code>joint</code>	Numeric variable. A discrete variable indicating pairs of candidates in double-member constituencies who filed their campaign spending returns as joint candidates. The variable is set to missing for all non-joint candidates. The variable has 808 non-missing observations and 403 distinct values.

<code>cand_id</code>	Numeric variable. Discrete variable that uniquely identifies an individual parliamentary candidate. The variable has 78,548 non-missing observations and 40,578 distinct values.
<code>name</code>	String variable. String variable reporting the name of the candidate. The format is Lastname, Fornames . When multiple candidates have the same name additional identifying information is included after a semi-colon. The variable has 78,548 non-missing observations and 38,153 distinct values.
<code>prefix</code>	String variable. String variable reporting titles placed before the candidate's name on the spending returns. The variable has 10,206 non-missing observations and 448 distinct values.
<code>suffix</code>	String variable. String variable reporting titles placed after the candidate's name on the spending returns. The variable has 6,202 non-missing observations and 1,218 distinct values.
<code>mp</code>	Numeric variable. A dummy variable indicating whether the candidate was a member of parliament at the time of the election (<code>==1</code>) or not a member of parliament (<code>==0</code>). The variable has 77,210 non-missing observations and 2 distinct values.
<code>inc</code>	Numeric variable. A dummy variable indicating whether the candidate is an incumbent (<code>==1</code>), or running as a challenger (<code>==0</code>). The variable has 78,548 non-missing observations and 2 distinct values.
<code>party</code>	String variable. String variable that reports the party affiliation of candidates. The variable has 78,548 non-missing observations and 30 distinct values.

<code>votes</code>	Numeric variable. Discrete variable that reports the total number of votes received by a candidate in a parliamentary election. The variable has 76,557 non-missing observations and 27,197 distinct values.
<code>vote_pct</code>	Numeric variable. Continuous variable that reports a candidate's percent of the votes cast in a parliamentary election. The variable has 76,557 non-missing observations and 76,375 distinct values.
<code>win</code>	Numeric variable. A dummy variable indicating whether the candidate won a parliamentary seat (<code>==1</code>) or lost (<code>==0</code>) The variable has 78,548 non-missing observations and 2 distinct values.
<code>spending_filed</code>	Numeric variable. Dummy variable indicating whether the candidate filed their campaign spending report (<code>==1</code>) or not (<code>==0</code>). The variable has 78,548 non-missing observations and 2 distinct values.
<code>spending_pct</code>	Numeric variable. Continuous variable that reports a candidate's percent of the total spending in a parliamentary election. The variable has 71,409 non-missing observations and 67,045 distinct values.
<code>spending_limit_predicted</code>	Numeric variable. Continous variable reporting the spending limit as calculated using the spending limit formula and the reported formula inputs. The amount is measured in british pounds in current prices. The variable has 73,606 non-missing observations and 7,009 distinct values.
<code>spending_limit_predicted_adj</code>	Numeric variable. Continous variable reporting the spending limit as calculated using the spending limit formula and the reported formula inputs. The amount is measured in

british pounds in constant prices (2020 prices). The variable has 73,606 non-missing observations and 12,740 distinct values.

`spending_limit_reported`

Numeric variable. Continuous variable reporting the spending limit as reported in the spending limit reports. The amount is measured in british pounds in current prices. The variable has 68,155 non-missing observations and 8,684 distinct values.

`spending_limit_reported_adj`

Numeric variable. Continuous variable reporting the spending limit as reported in the spending limit reports. The amount is measured in british pounds in constant prices (2020 prices). The variable has 68,155 non-missing observations and 13,274 distinct values.

`spending_limited`

Numeric variable. Continuous variable measuring candidates' aggregate expenditure that counts toward the legally permitted level of spending. The amount is measured in british pounds in current prices. The variable has 71,395 non-missing observations and 35,847 distinct values.

`spending_limited_adj`

Numeric variable. Continuous variable measuring candidates' aggregate expenditure that counts toward the legally permitted level of spending. The amount is measured in british pounds in constant prices (2020 prices). The variable has 71,395 non-missing observations and 51,220 distinct values.

`spending_abovelimit`

Numeric variable. Dummy variable that reports whether the candidate spends more than the legally permitted amount on their electoral campaign (`==1`), or not (`==0`). The

variable has 70,959 non-missing observations and 2 distinct values.

`spending_total` Numeric variable. Continuous variable measuring candidates' total expenditure. The amount is measured in british pounds in current prices. The variable has 71,827 non-missing observations and 36,573 distinct values.

`spending_total_adj` Numeric variable. Continuous variable measuring candidates' total expenditure. The amount is measured in british pounds in constant prices (2020 prices). The variable has 71,827 non-missing observations and 52,321 distinct values.

`spending_ads` Numeric variable. Continuous variable measuring the amount spent on advertisement and printing. The amount is measured in british pounds in current prices. The variable has 68,164 non-missing observations and 32,040 distinct values.

`spending_ads_adj` Numeric variable. Continuous variable measuring the amount spent on advertisement and printing. The amount is measured in british pounds in constant prices (2020 prices). The variable has 68,164 non-missing observations and 44,836 distinct values.

`spending_managers` Numeric variable. Continuous variable measuring the amount spent on campaign managers. This includes money spent on agents and subagents. The amount is measured in british pounds in current prices. The variable has 53,330 non-missing observations and 5,684 distinct values.

`spending_managers_adj` Numeric variable. Continuous variable measuring the amount spent on campaign managers. This includes money spent on agents and subagents. The amount is measured in british

pounds in constant prices (2020 prices). The variable has 53,330 non-missing observations and 10,459 distinct values.

`spending_staff` Numeric variable. Continuous variable measuring the amount spent on staff and clerks. The amount is measured in british pounds in current prices. The variable has 53,330 non-missing observations and 7,602 distinct values.

`spending_staff_adj` Numeric variable. Continuous variable measuring the amount spent on staff and clerks. The amount is measured in british pounds in constant prices (2020 prices). The variable has 53,330 non-missing observations and 12,658 distinct values.

`spending_allemployees` Numeric variable. Continuous variable measuring the aggregate amount spent on paying all employees, managers and staff, working on the campaign. The amount is measured in british pounds in current prices. The variable has 67,841 non-missing observations and 11,856 distinct values.

`spending_allemployees_adj` Numeric variable. Continuous variable measuring the amount spent on campaign headquarters and other rooms during the electoral campaign. The amount is measured in british pounds in constant prices (2020 prices). The variable has 67,841 non-missing observations and 18,926 distinct values.

`spending_facilities` Numeric variable. Continuous variable measuring the amount spent on campaign headquarters and other rooms during the electoral campaign. The amount is measured in british pounds in current prices. The variable has 68,164 non-missing observations and 10,717 distinct values.

- `spending_facilities_adj` Numeric variable. Continuous variable measuring the amount spent on campaign headquarters and other rooms during the electoral campaign. The amount is measured in british pounds in constant prices (2020 prices). The variable has 68,164 non-missing observations and 17,267 distinct values.
- `spending_meetings` Numeric variable. Continuous variable measuring the amount spent on organizing public meetings. The amount is measured in british pounds in current prices. The variable has 68,164 non-missing observations and 6,911 distinct values.
- `spending_meetings_adj` Numeric variable. Continuous variable measuring the amount spent on organizing public meetings. The amount is measured in british pounds in constant prices (2020 prices). The variable has 68,164 non-missing observations and 12,829 distinct values.
- `spending_misc` Numeric variable. Continuous variable measuring the amount spent on miscellaneous matters, i.e. expenses that are include in the other spending categories. The amount is measured in british pounds in current prices. The variable has 68,164 non-missing observations and 16,150 distinct values.
- `spending_misc_adj` Numeric variable. Continuous variable measuring the amount spent on miscellaneous matters, i.e. expenses that are include in the other spending categories. The amount is measured in british pounds in constant prices (2020 prices). The variable has 68,164 non-missing observations and 22,296 distinct values.
- `spending_personal` Numeric variable. Continous variable measuring candidates' personal campaign expenses. The amount is measured in

british pounds in current prices. The variable has 71,718 non-missing observations and 8,084 distinct values.

`spending_personal_adj`

Numeric variable. Continuous variable measuring candidates' personal campaign expenses. The amount is measured in british pounds in constant prices (2020 prices). The variable has 71,718 non-missing observations and 13,451 distinct values.

`cpi_multiplier`

Numeric variable. Consumer-price index multiplier. The variable has 78,548 non-missing observations and 130 distinct values.

`eggerts_spirling_constituency_id`

Numeric variable. Numeric variable that reports the constituency identifier used by Eggerts and Spirling. The variable has 78,547 non-missing observations and 1,907 distinct values.

`eggerts_spirling_id`

Numeric variable. Numeric variable that reports the observation identifier used by Eggerts and Spirling. The variable has 67,655 non-missing observations and 67,655 distinct values.

`next_election_in_constituency`

Numeric variable. Numeric variable that reports the date of the next election in the constituency. The variable has 62,818 non-missing observations and 1,637 distinct values.

`prev_election_in_constituency`

Numeric variable. Numeric variable that reports the date of the previous election in the constituency. The variable has 63,189 non-missing observations and 1,365 distinct values.

<code>county</code>	Numeric variable. Discrete variable that reports whether a constituency is classified as either county (<code>==1</code>) or borough (<code>==0</code>). University constituencies are coded as missing. The variable has 78,208 non-missing observations and 2 distinct values.
<code>notes_spending_limit</code>	String variable. String variable that describes issues with the reported spending limit. The variable has 3,075 non-missing observations and 2 distinct values.
<code>notes_spending</code>	String variable. String variable that describes issues with the reported spending. The variable has 150 non-missing observations and 3 distinct values.

Table 2.1: Descriptive Statistics

	Mean	Standard Dev.	Min.	Max.	Obs.
nation	1.66	1.30	1.00	5.00	78,548
region	7.40	3.57	1.00	13.00	78,548
year	1966.52	40.28	1885.00	2019.00	78,548
by_election	0.06	0.25	0.00	1.00	78,548
seats_up	1.02	0.14	1.00	3.00	78,548
seats_total	1.02	0.14	1.00	3.00	78,548
electors	53964.68	23096.57	0.00	1.7e+05	75,772
turnout	0.72	0.10	0.01	1.47	74,261
candidates	4.05	2.07	1.00	26.00	78,548
polling_districts	38.84	27.18	1.00	229.00	48,971
polling_stations	63.39	32.04	1.00	998.00	48,835
joint	201.88	116.29	1.00	403.00	808
mp	0.24	0.43	0.00	1.00	77,210
inc	0.20	0.40	0.00	1.00	78,548
votes	10580.49	9372.81	1.00	75205.00	76,557
vote_pct	29.84	21.33	0.00	100.00	76,557
win	0.32	0.47	0.00	1.00	78,548
spending_files	0.91	0.28	0.00	1.00	78,548
spending_pct	30.69	20.84	0.00	100.00	71,409
spending_limit_predicted	6000.81	9358.03	0.00	1.0e+05	73,606
spending_limit_predicted_adj	38165.47	39819.57	0.00	4.5e+05	73,606
spending_limit_reported	5258.09	7681.21	24.00	5.7e+05	68,155
spending_limit_reported_adj	38669.30	40453.97	219.60	6.3e+05	68,155
spending_limited	2134.83	3410.03	0.00	98697.68	71,395
spending_limited_adj	21156.09	29721.04	0.00	4.5e+05	71,395
spending_abovelimit	0.00	0.05	0.00	1.00	70,959
spending_total	2190.75	3432.90	0.00	98697.68	71,827
spending_total_adj	22696.70	31770.98	0.00	4.5e+05	71,827
spending_ads	1602.31	2563.59	0.00	76557.93	68,164
spending_ads_adj	12297.11	13659.26	0.00	2.3e+05	68,164
spending_managers	91.34	153.63	0.00	4034.38	53,330
spending_managers_adj	4525.74	10190.07	0.00	1.0e+05	53,330
spending_staff	53.78	87.80	0.00	4937.91	53,330
spending_staff_adj	3208.15	5939.77	0.00	1.5e+05	53,330
spending_allemployees	186.69	757.13	0.00	52977.00	67,841
spending_allemployees_adj	6163.91	13231.77	0.00	2.0e+05	67,841
spending_facilities	118.13	479.90	0.00	36339.37	68,164
spending_facilities_adj	1094.57	1828.04	0.00	48892.40	68,164
spending_meetings	25.76	62.56	0.00	3373.00	68,164
spending_meetings_adj	841.90	1537.20	0.00	26500.84	68,164
spending_misc	107.78	255.59	0.00	8792.92	68,164
spending_misc_adj	1890.73	3222.08	0.00	44315.41	68,164
spending_personal	60.59	192.37	0.00	10588.94	71,718
spending_personal_adj	1273.65	2718.31	0.00	84052.64	71,718
cpi_multiplier	33.85	43.31	1.03	133.63	78,548
next_election_in_constituency	1918.62	14949.74	-2.7e+04	21895.00	62,818
prev_election_in_constituency	665.24	14671.63	-2.7e+04	20978.00	63,189
county	0.50	0.50	0.00	1.00	78,208

Chapter 3

Citation

Please cite the following paper when using the dataset:

```
@article{fournaies20limits,  
title={How Do Campaign Spending Limits Affect Elections? Evidence from United Kingdom 1885–2019},  
author={Alexander Fournaies},  
journal={American Political Science Review},  
volume={},  
number={},  
pages={},  
year={},  
publisher={Cambridge University Press}  
}
```

Bibliography

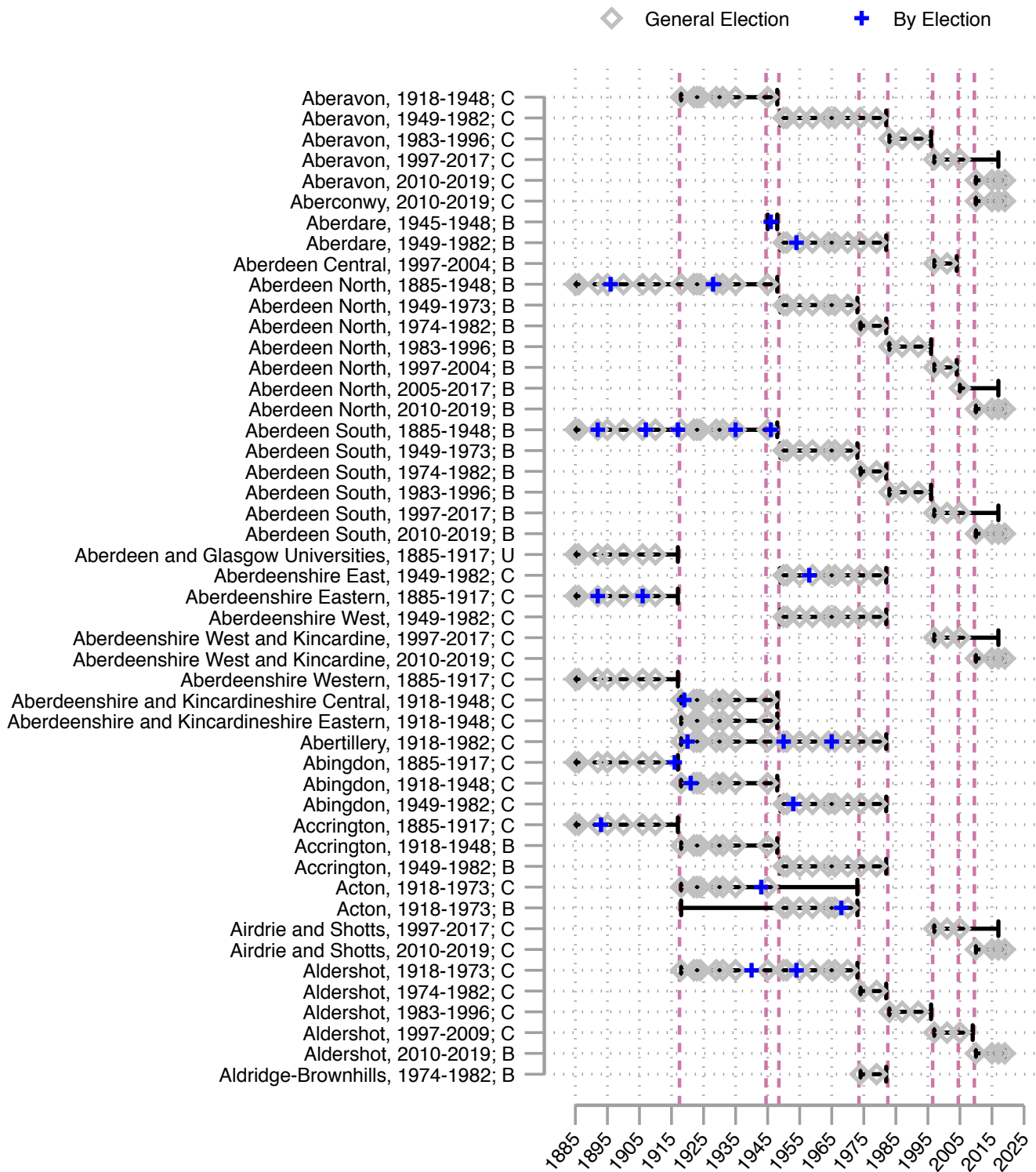
Eggers, Andrew C, and Arthur Spirling. 2014. “Electoral Security as a Determinant of Legislator Activity, 1832–1918: New Data and Methods for Analyzing British Political Development.” *Legislative Studies Quarterly* 39(4): 593–620.

Johnston, Ron, Charles Pattie, and Todd Hartman. 2019. “Spending, canvassing and electoral success in marginal constituencies: England and Scotland in 2017.” *Journal of Elections, Public Opinion and Parties* pp. 1–18.

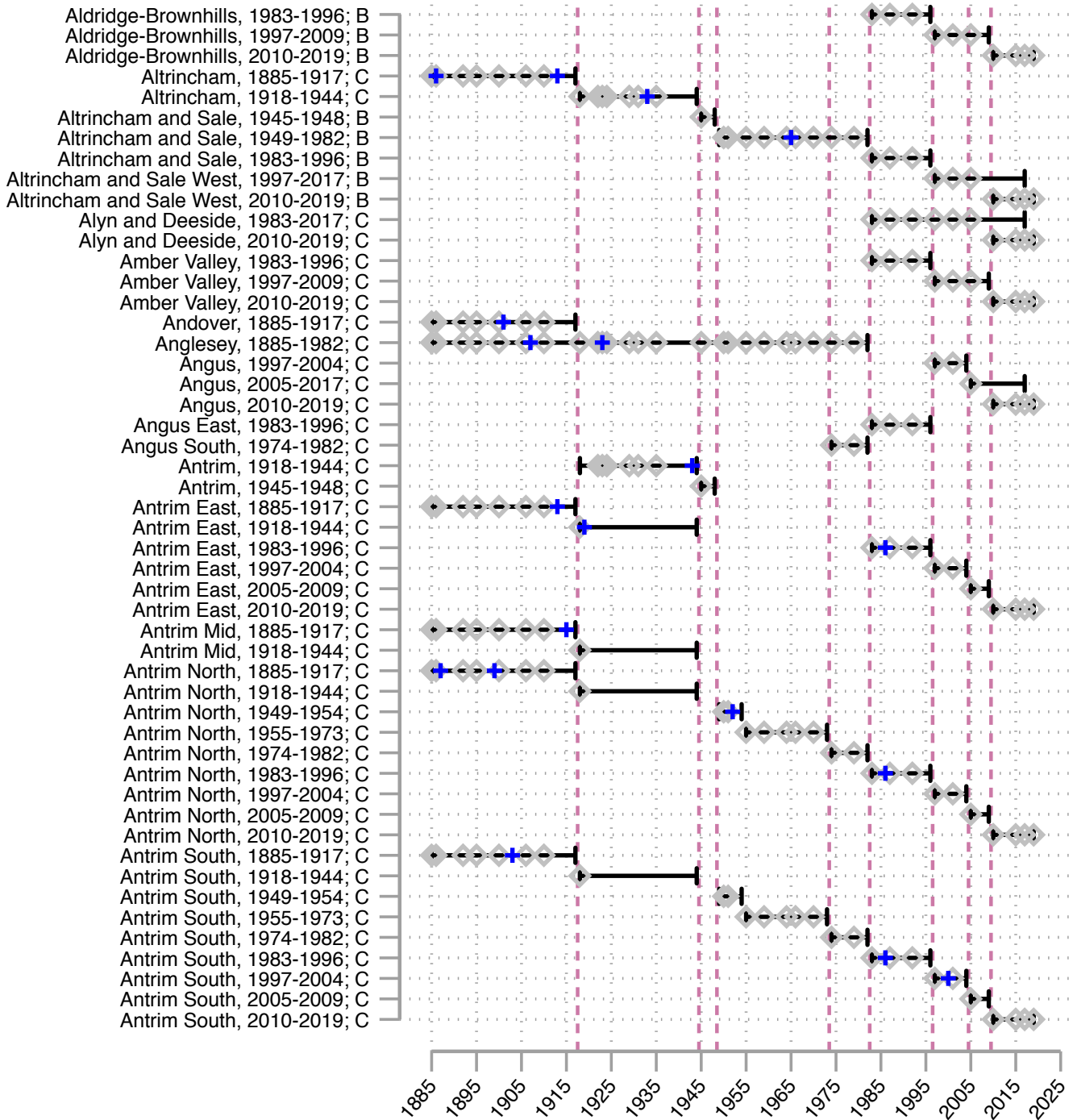
Chapter 4

Appendix

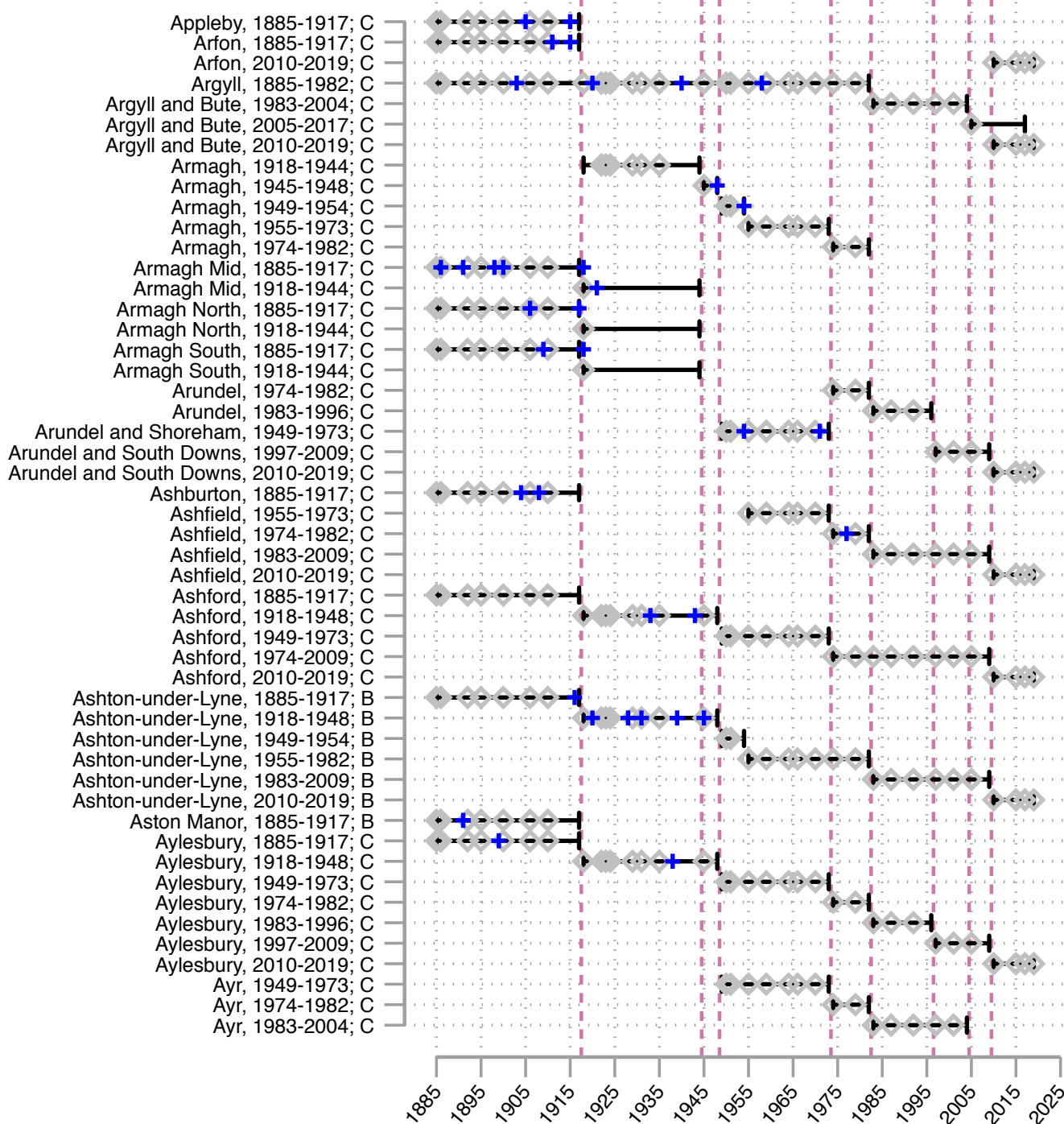
Contents



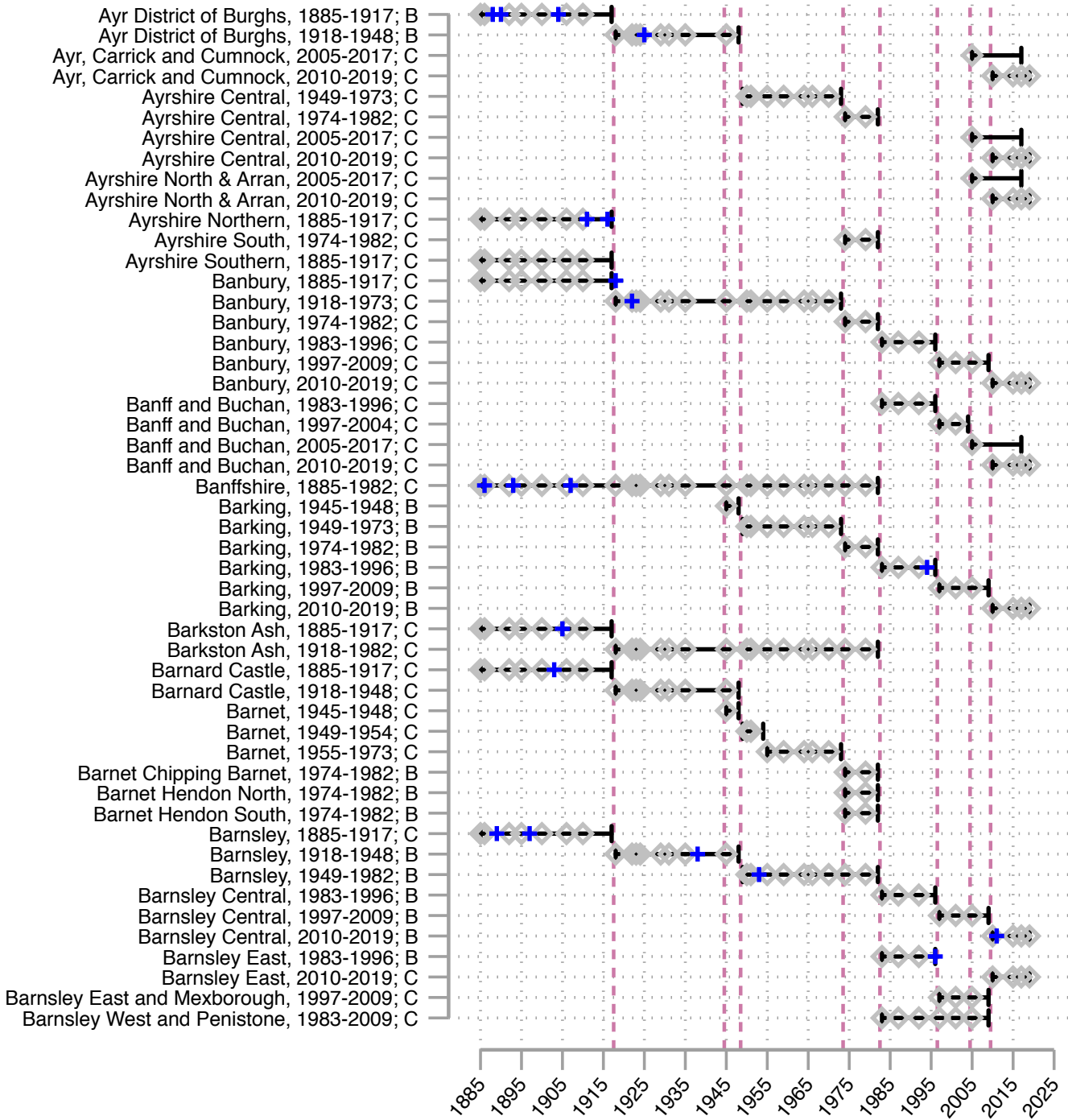
◇ General Election + By Election



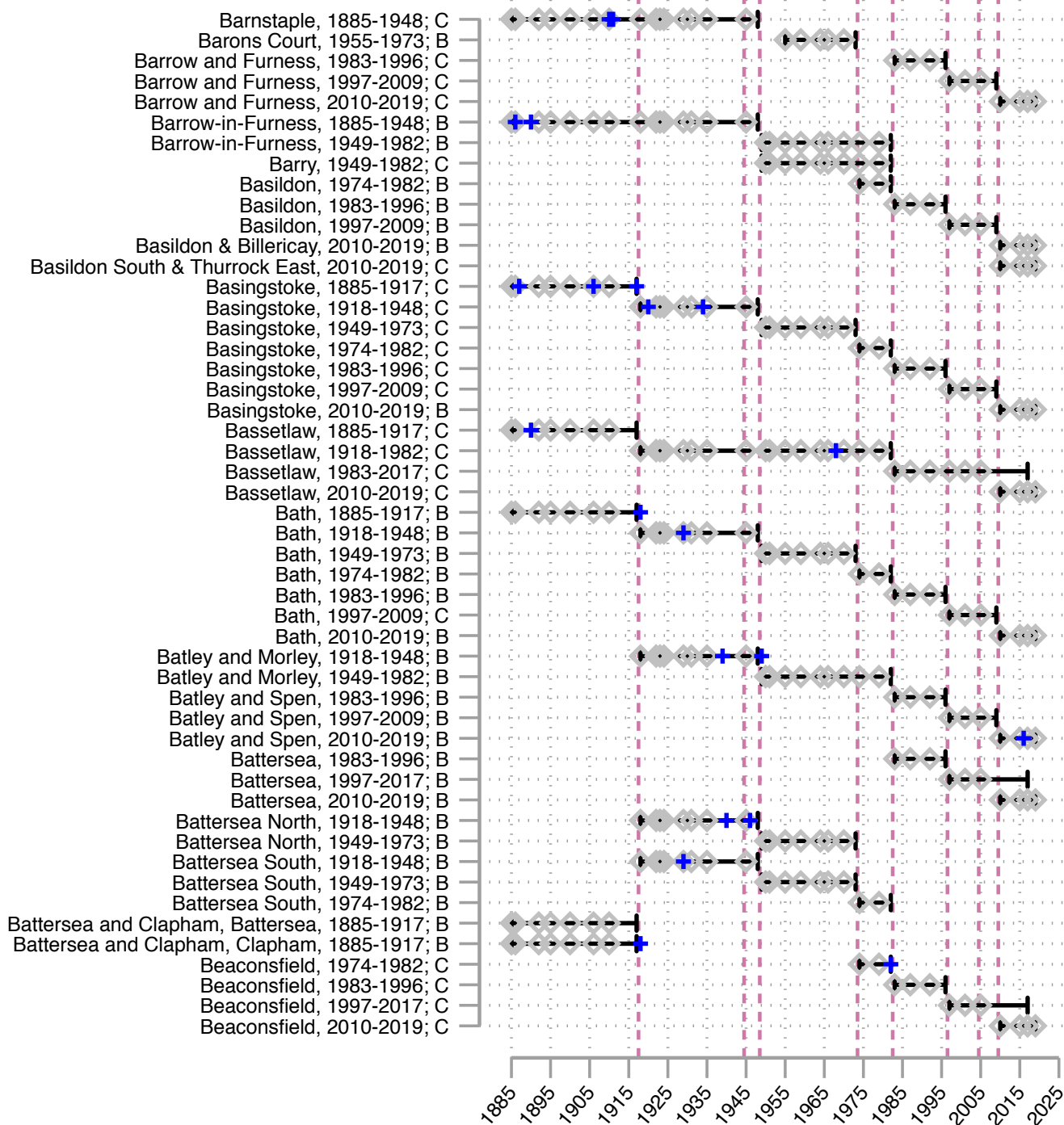
◇ General Election + By Election



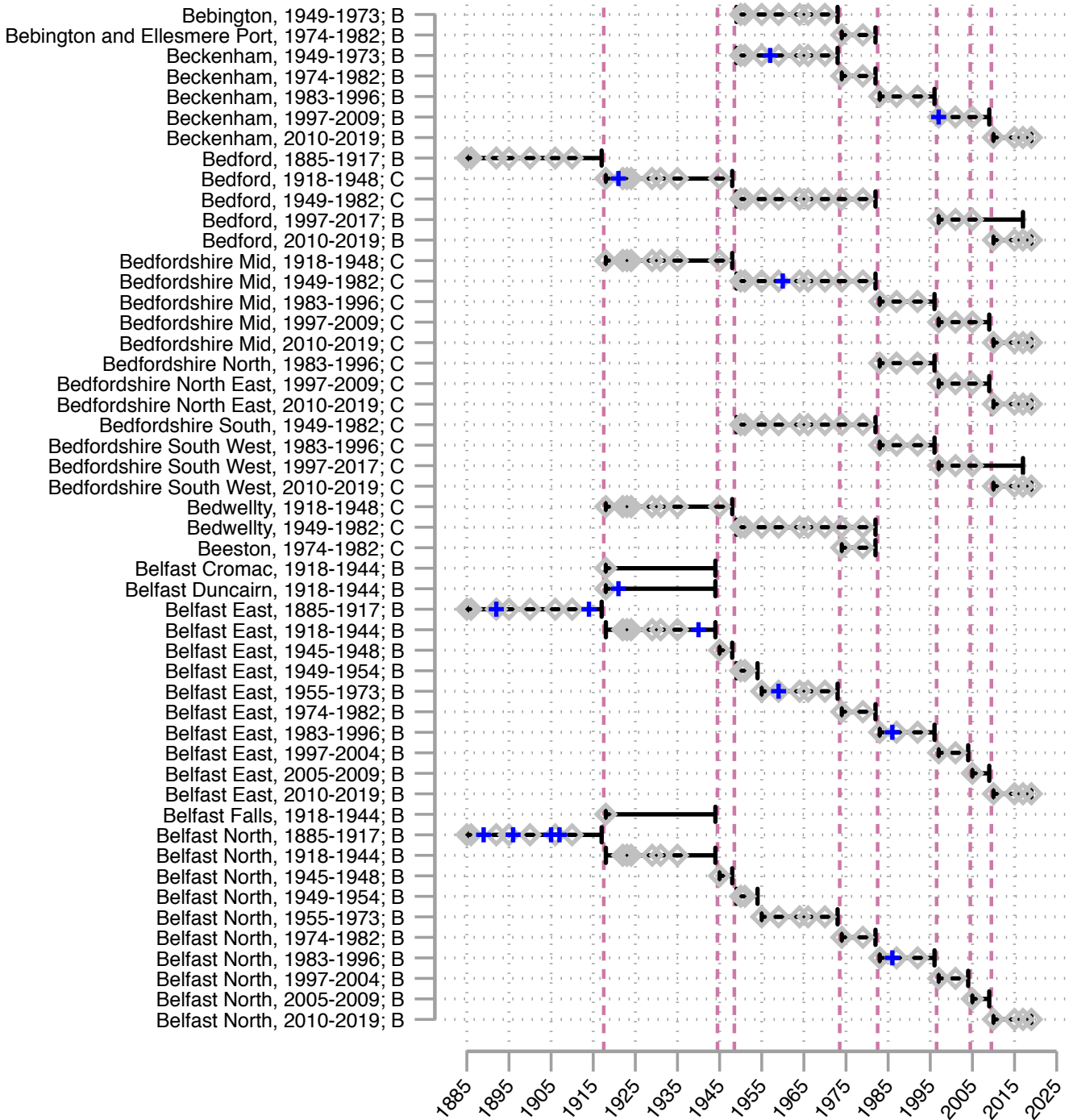
◇ General Election + By Election



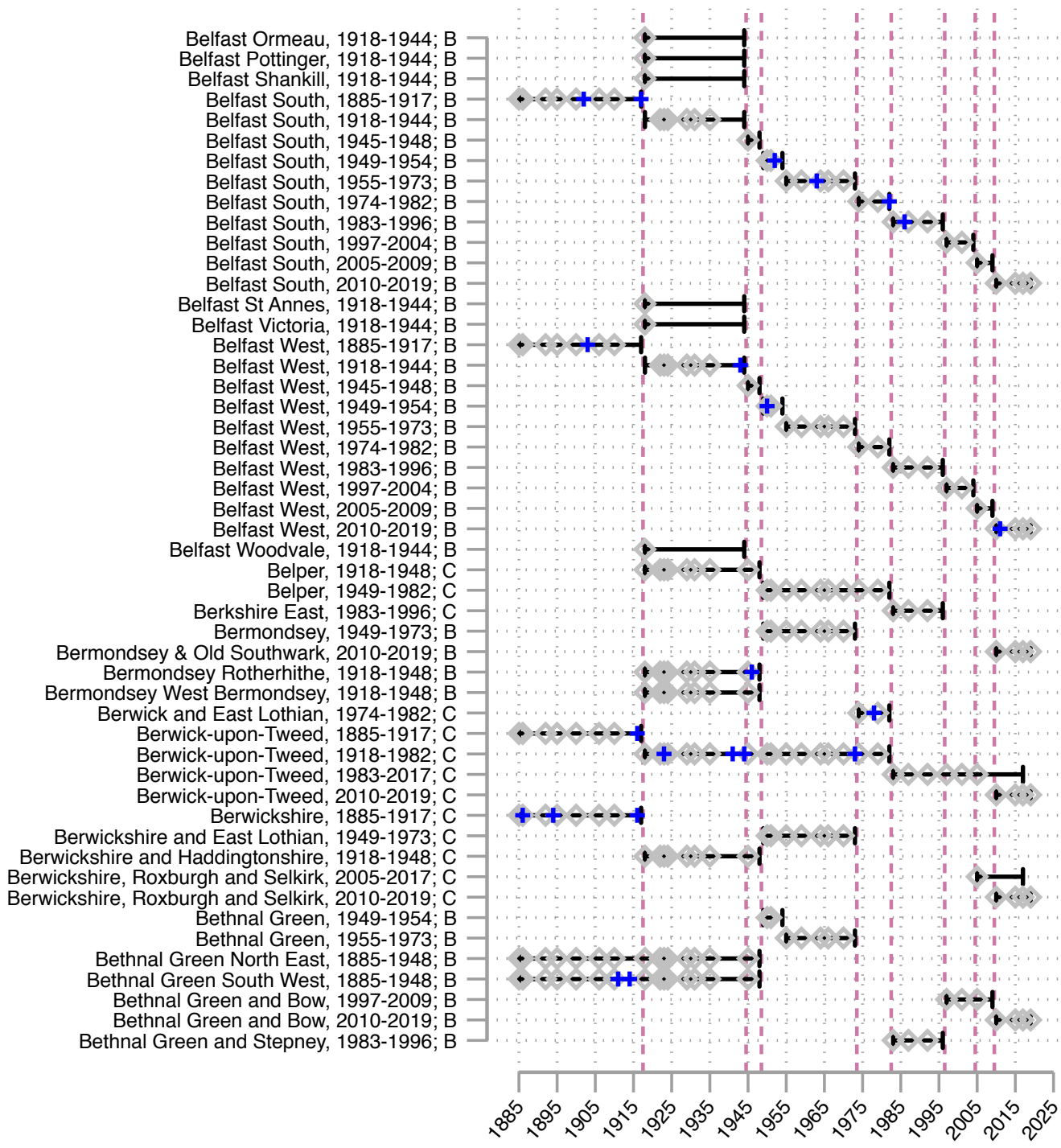
◇ General Election + By Election



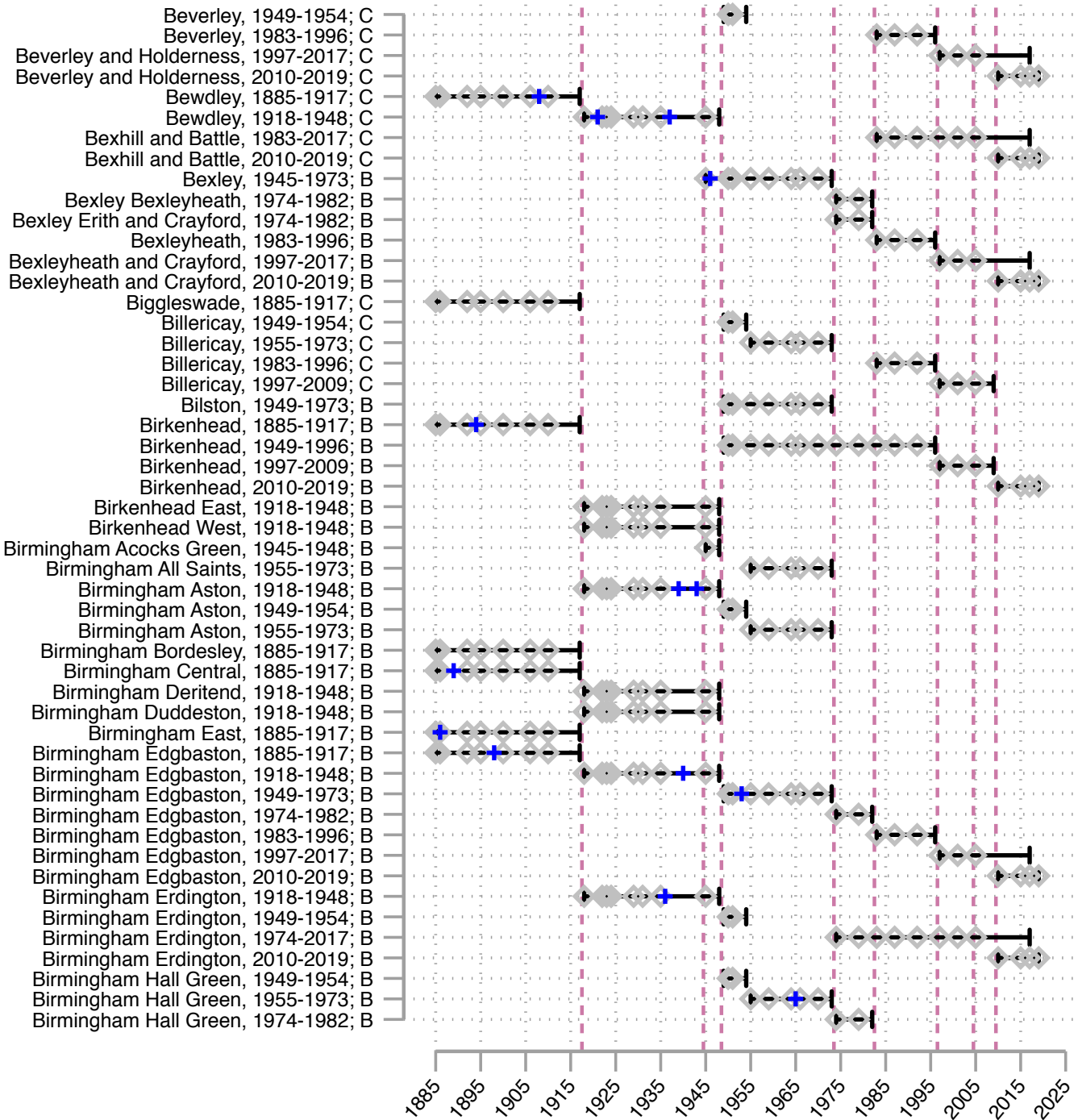
◇ General Election + By Election



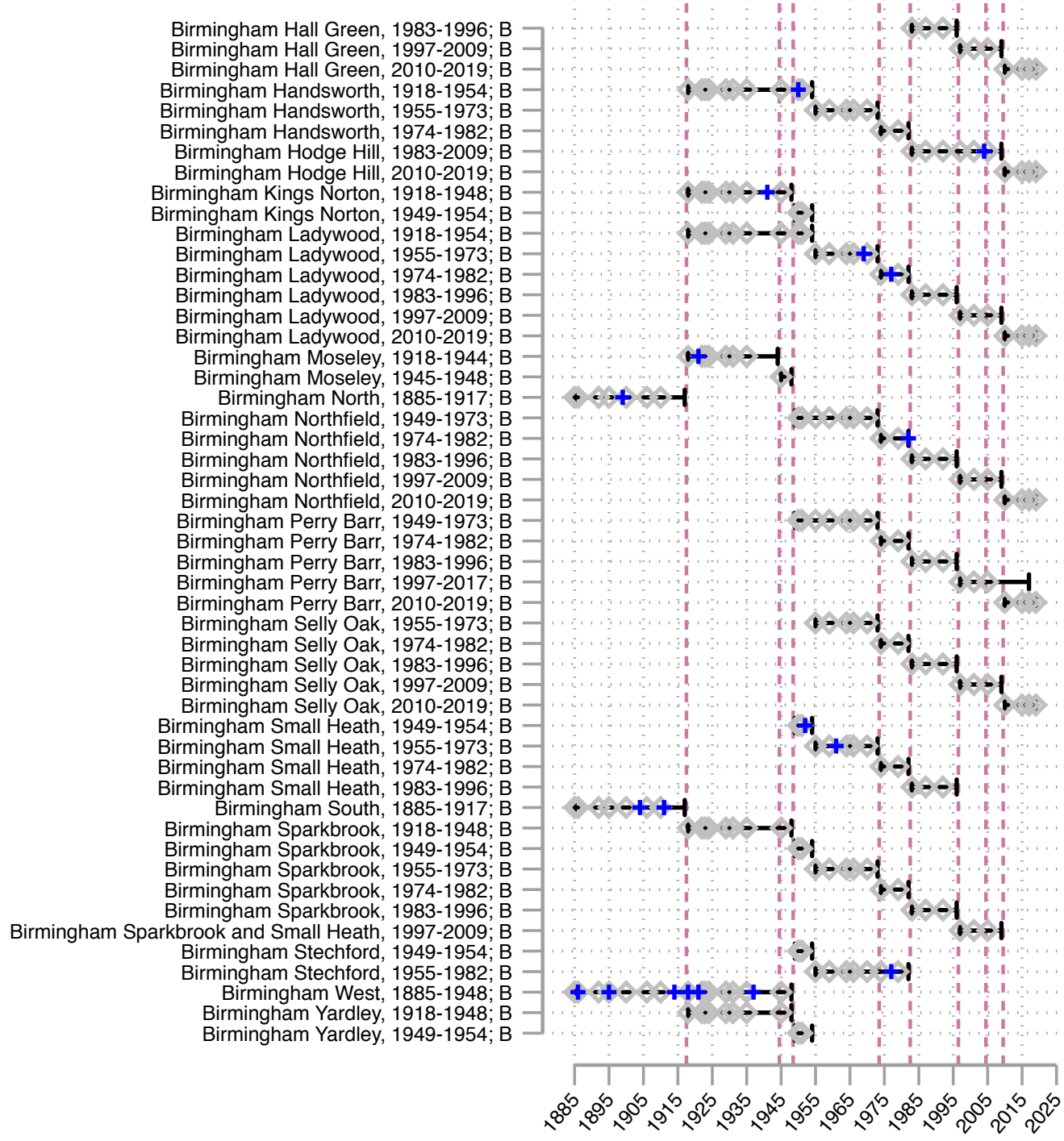
◇ General Election + By Election



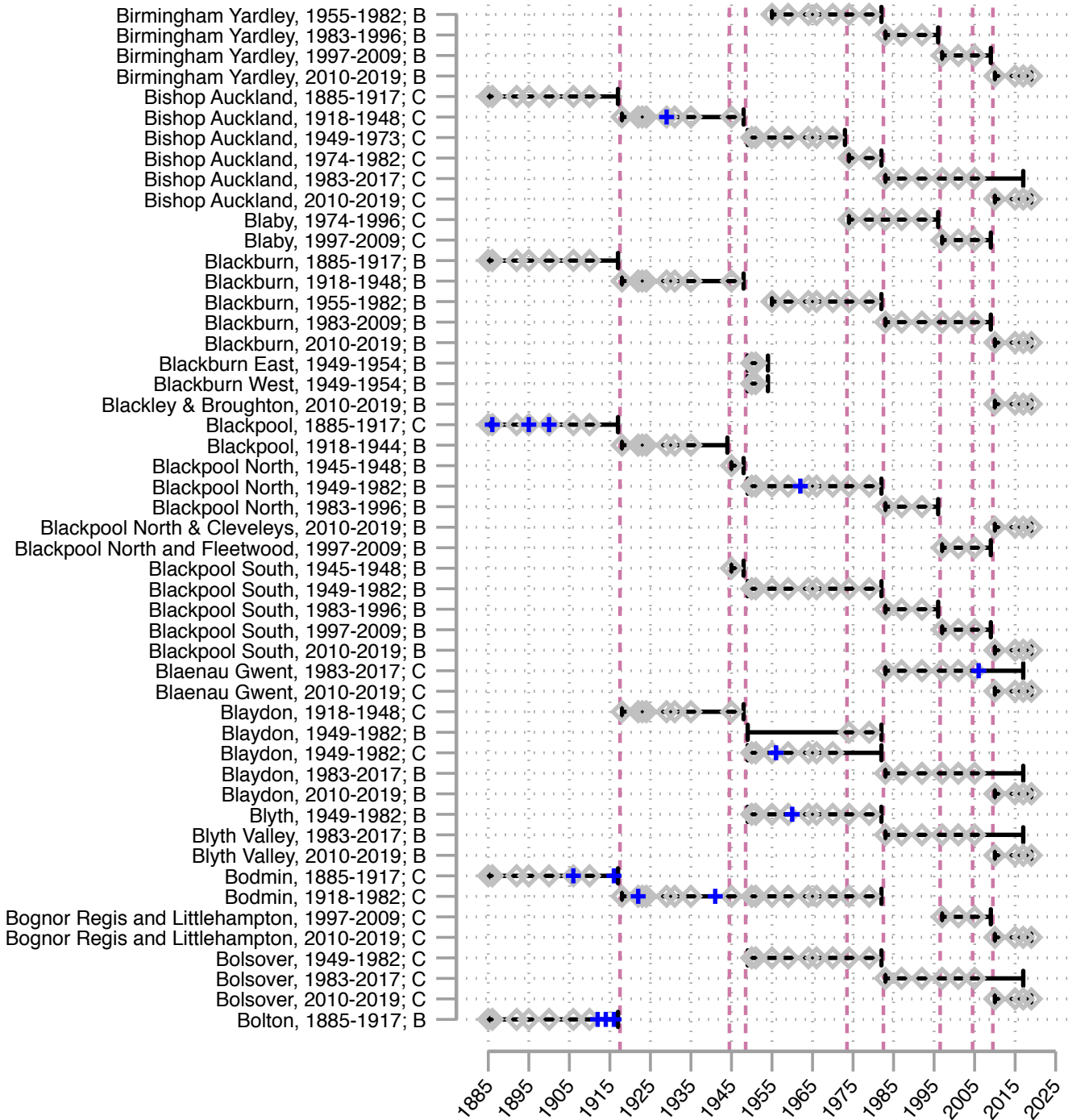
◇ General Election + By Election



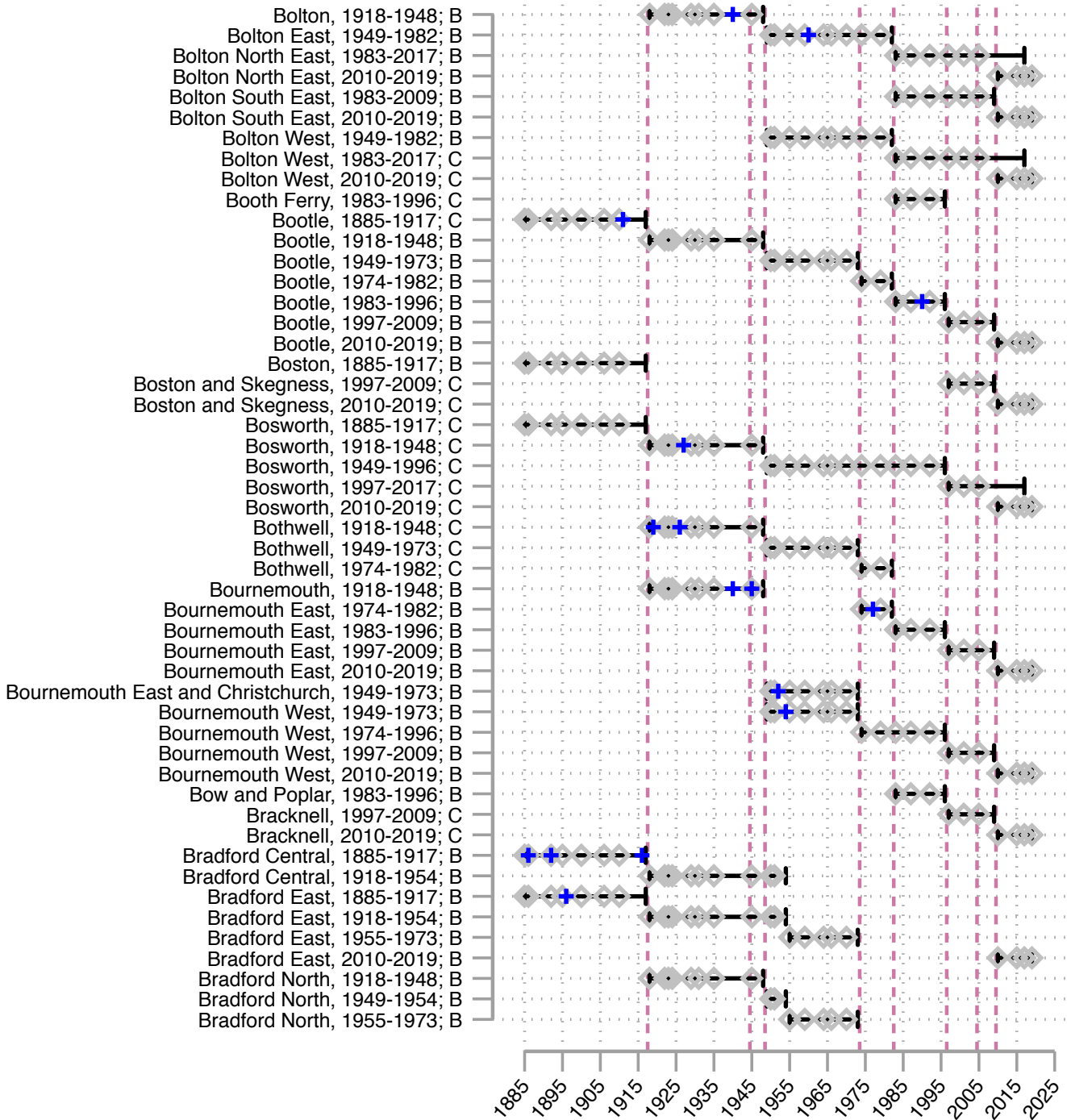
◇ General Election + By Election



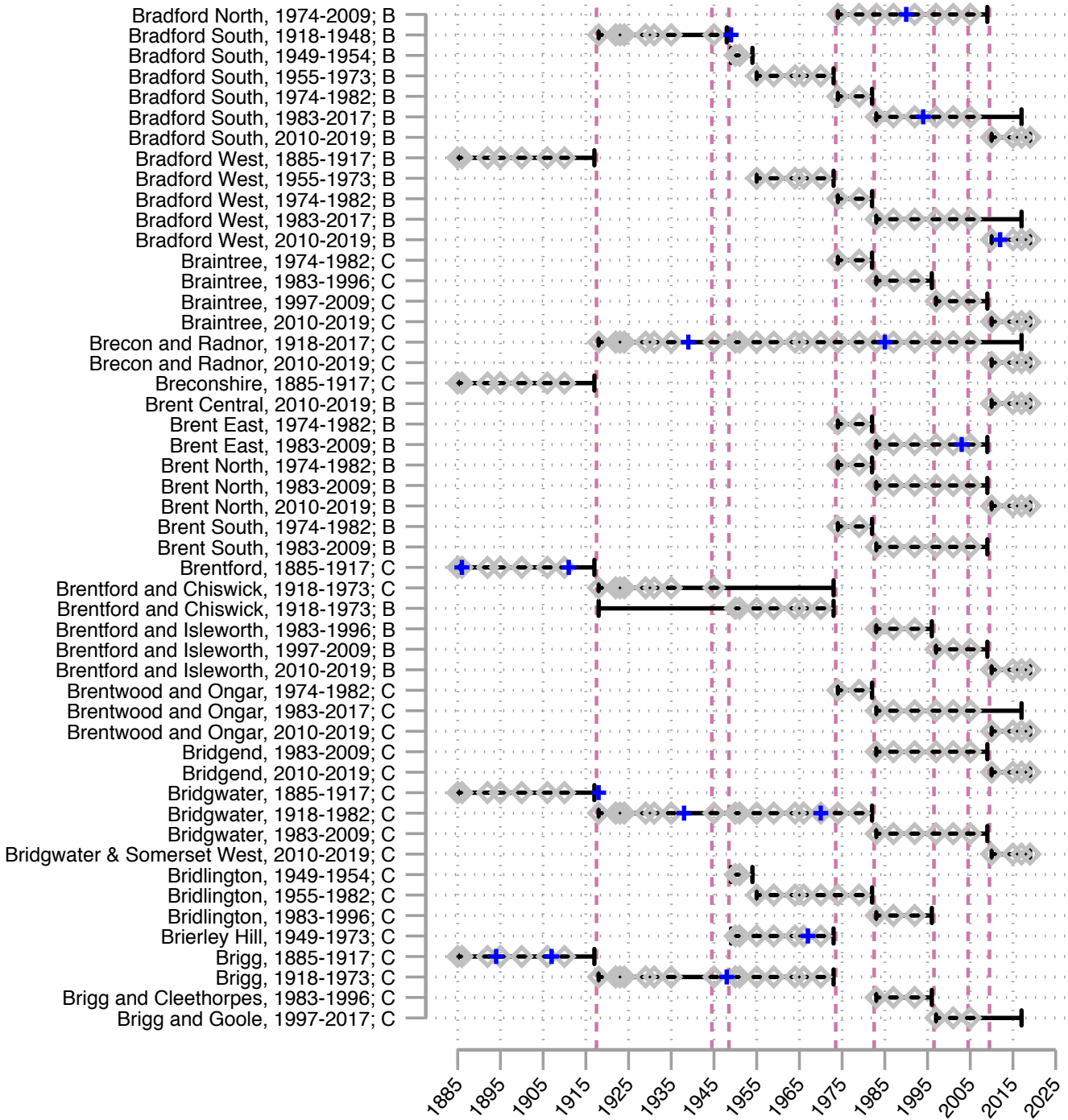
◇ General Election + By Election



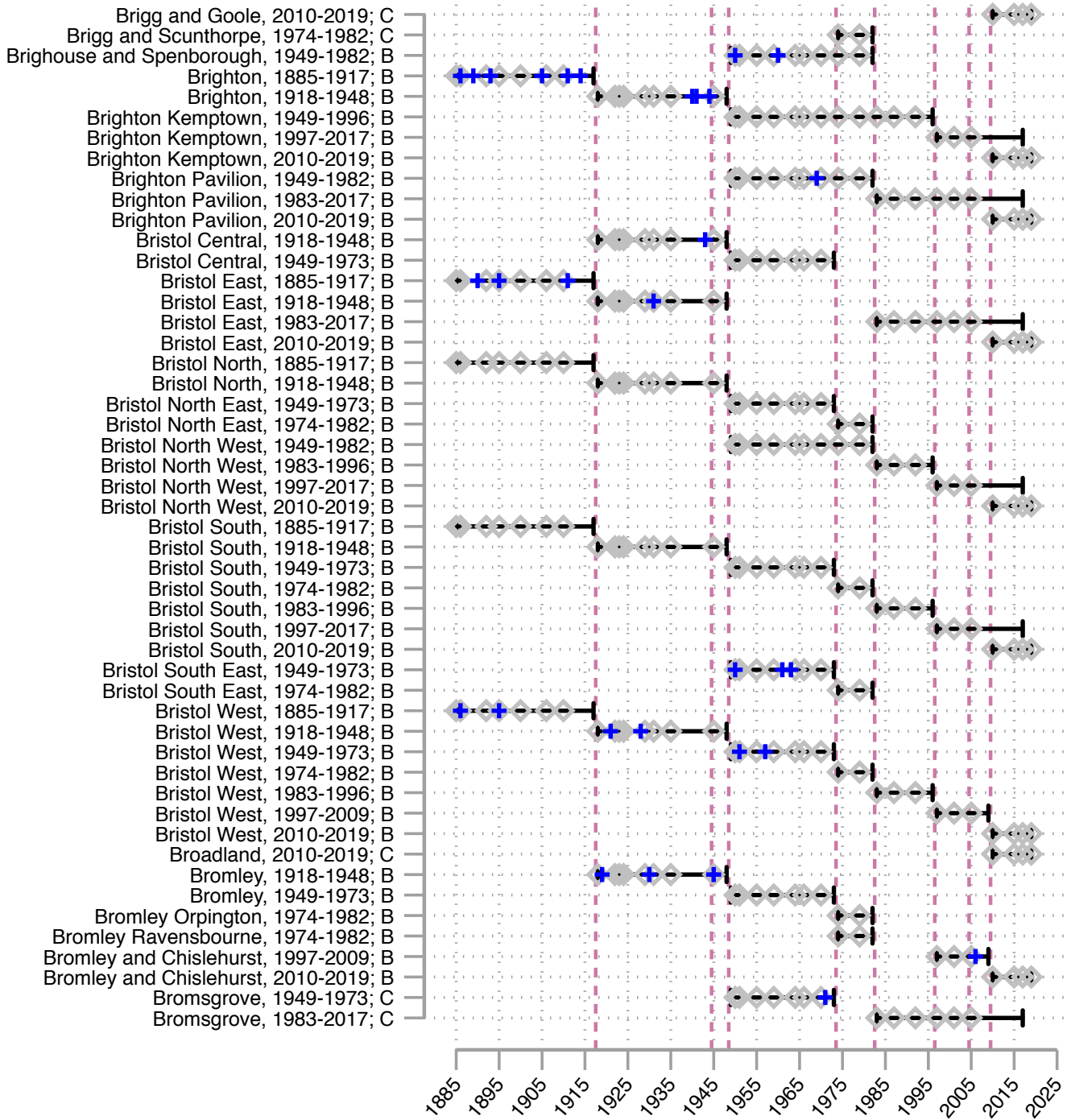
◇ General Election + By Election



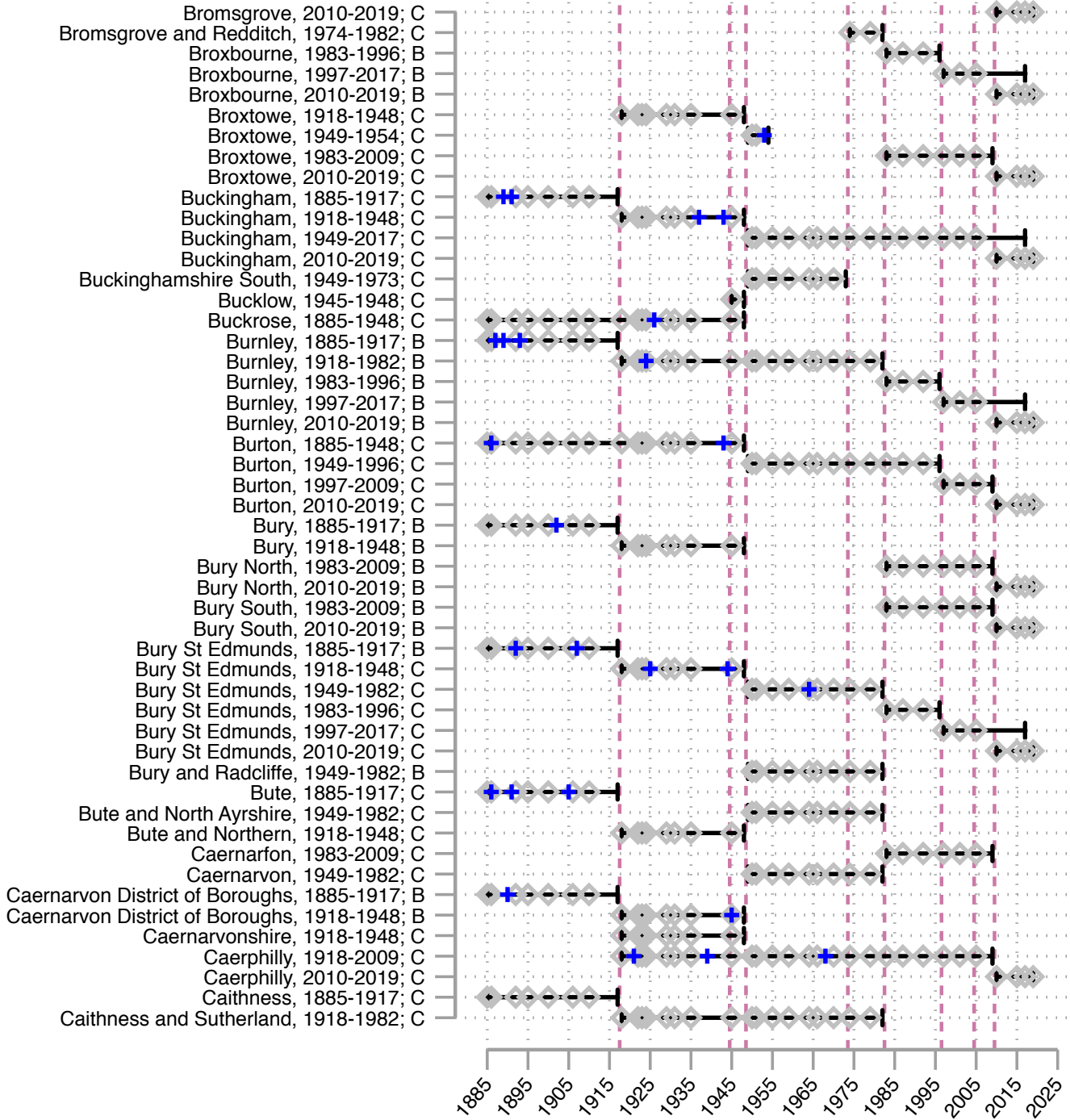
◇ General Election + By Election

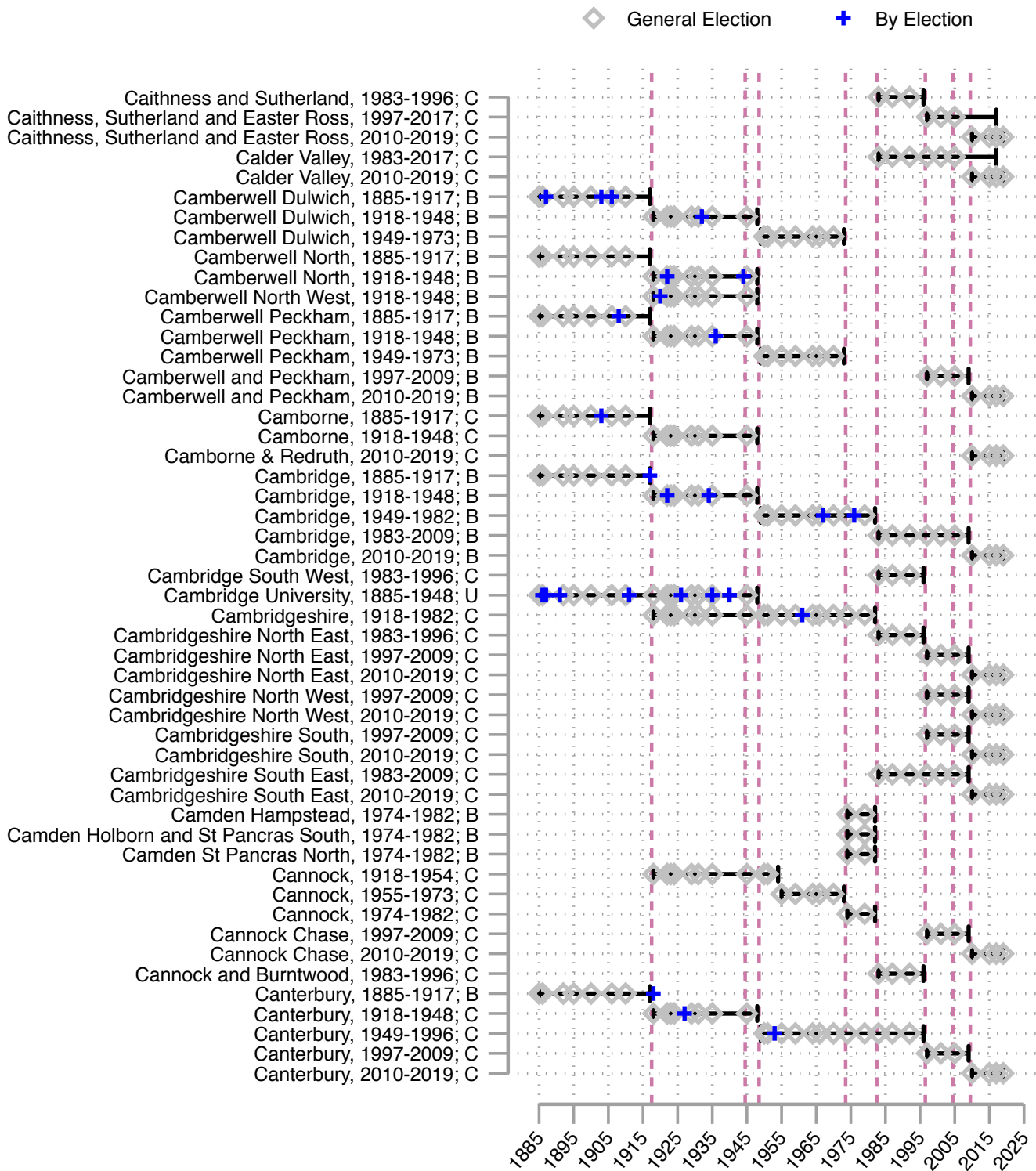


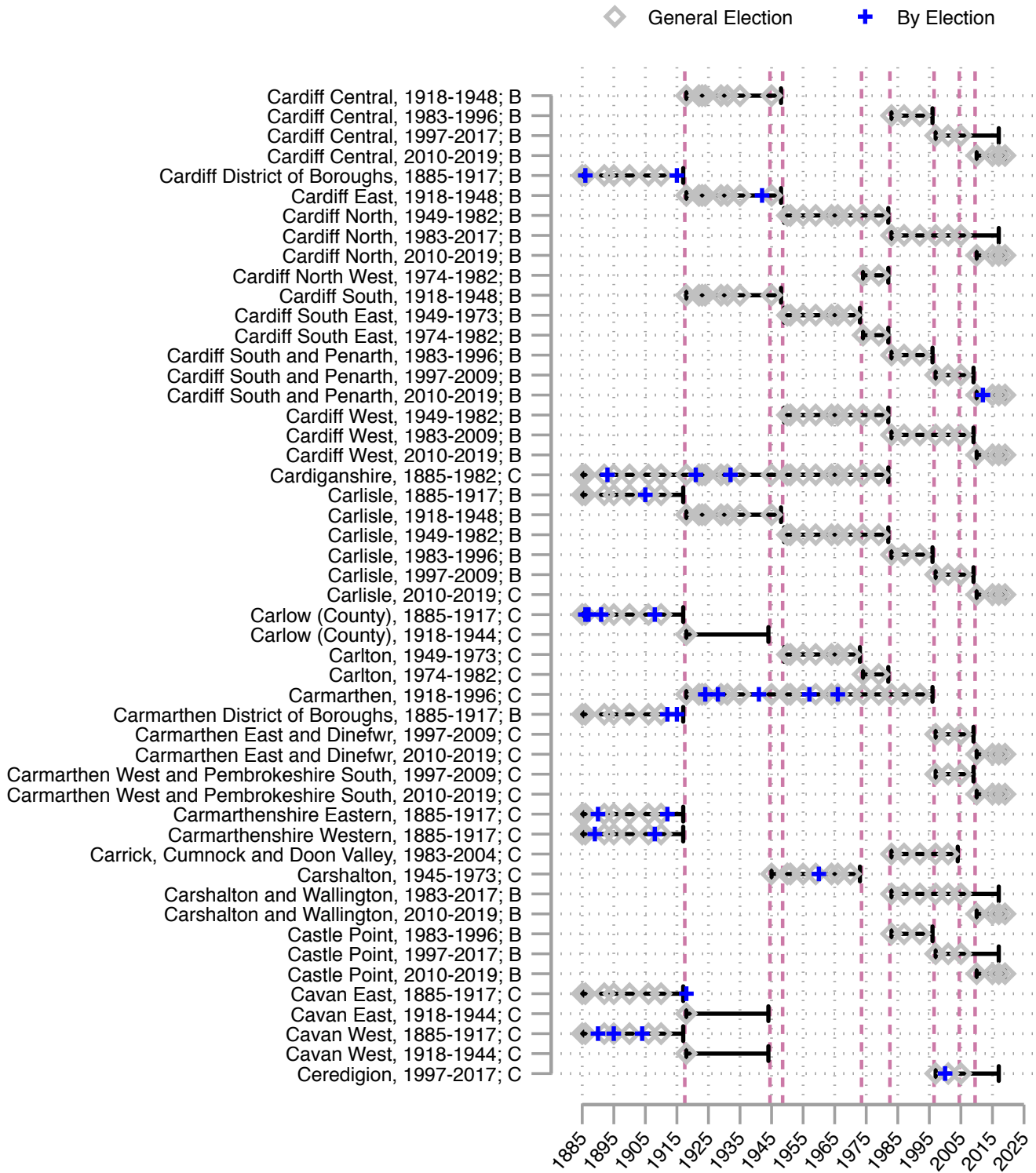
◇ General Election + By Election



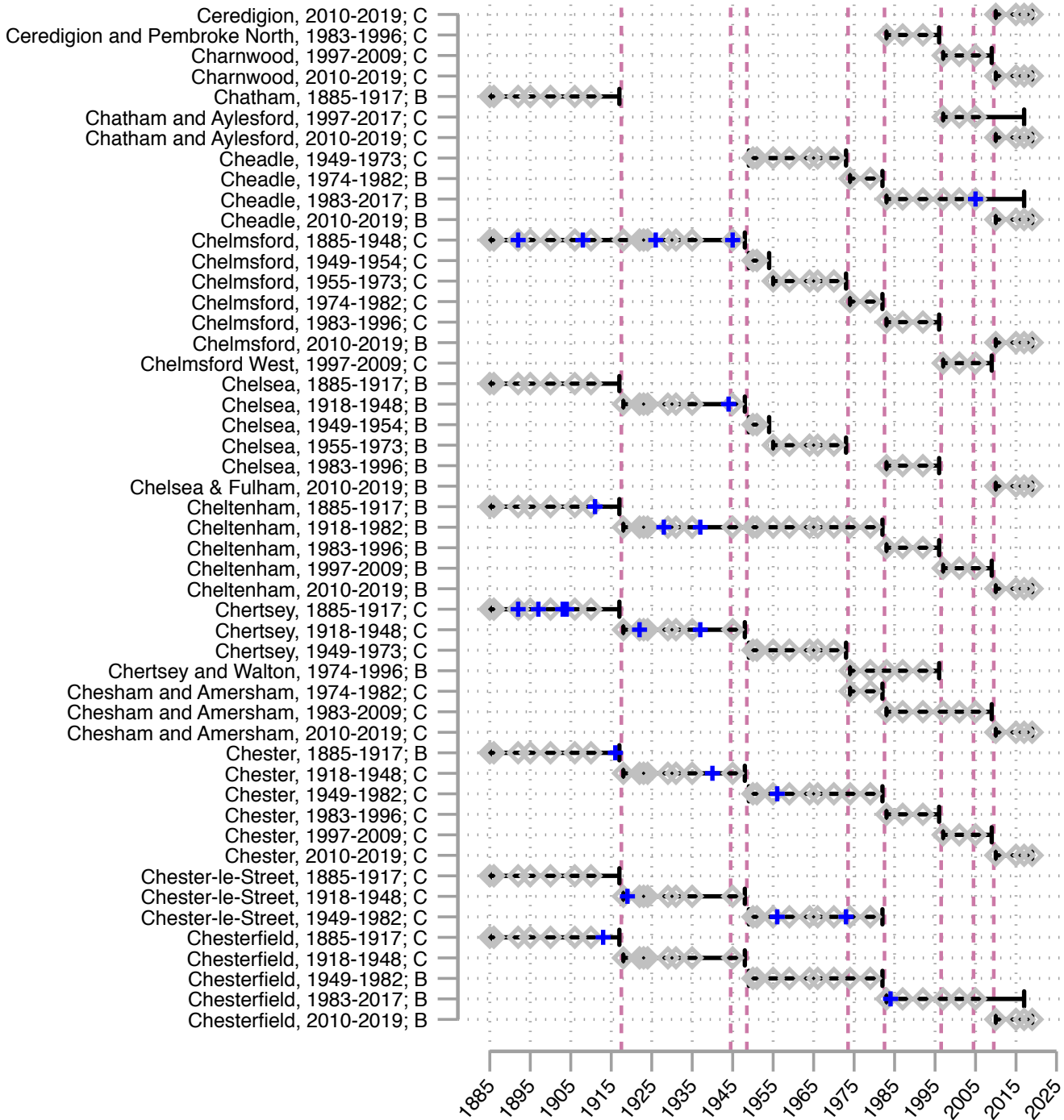
◇ General Election + By Election

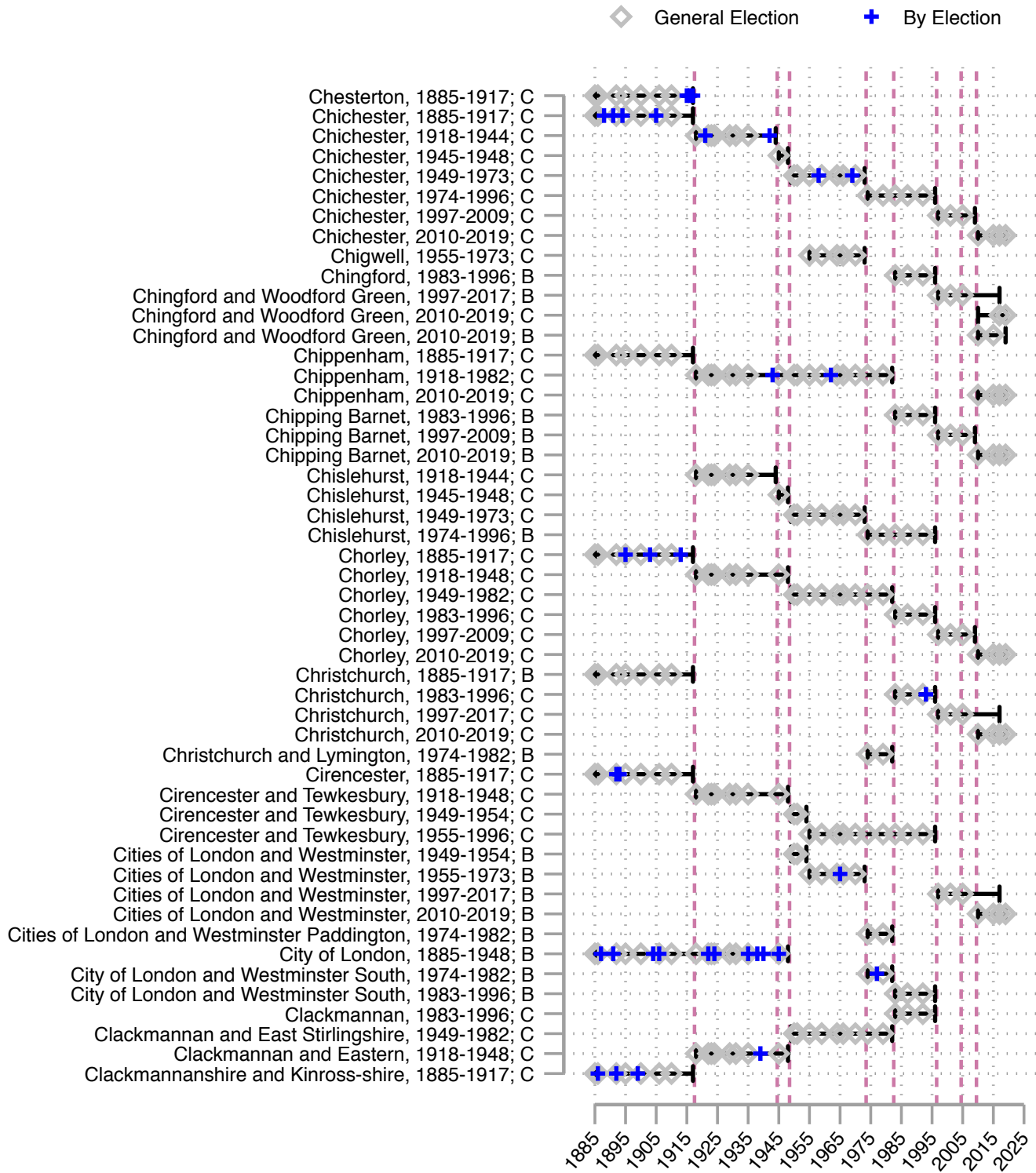




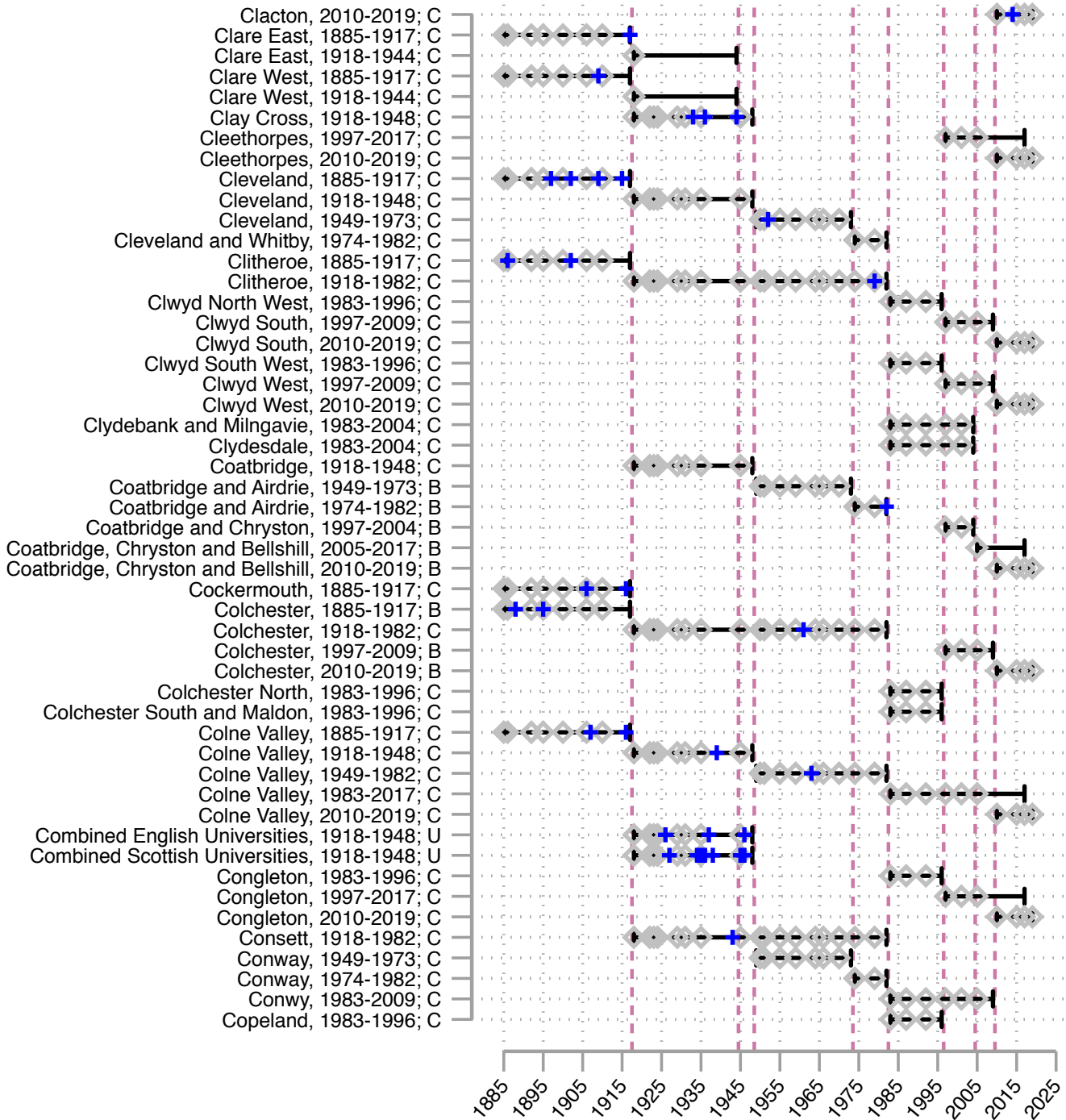


◇ General Election + By Election

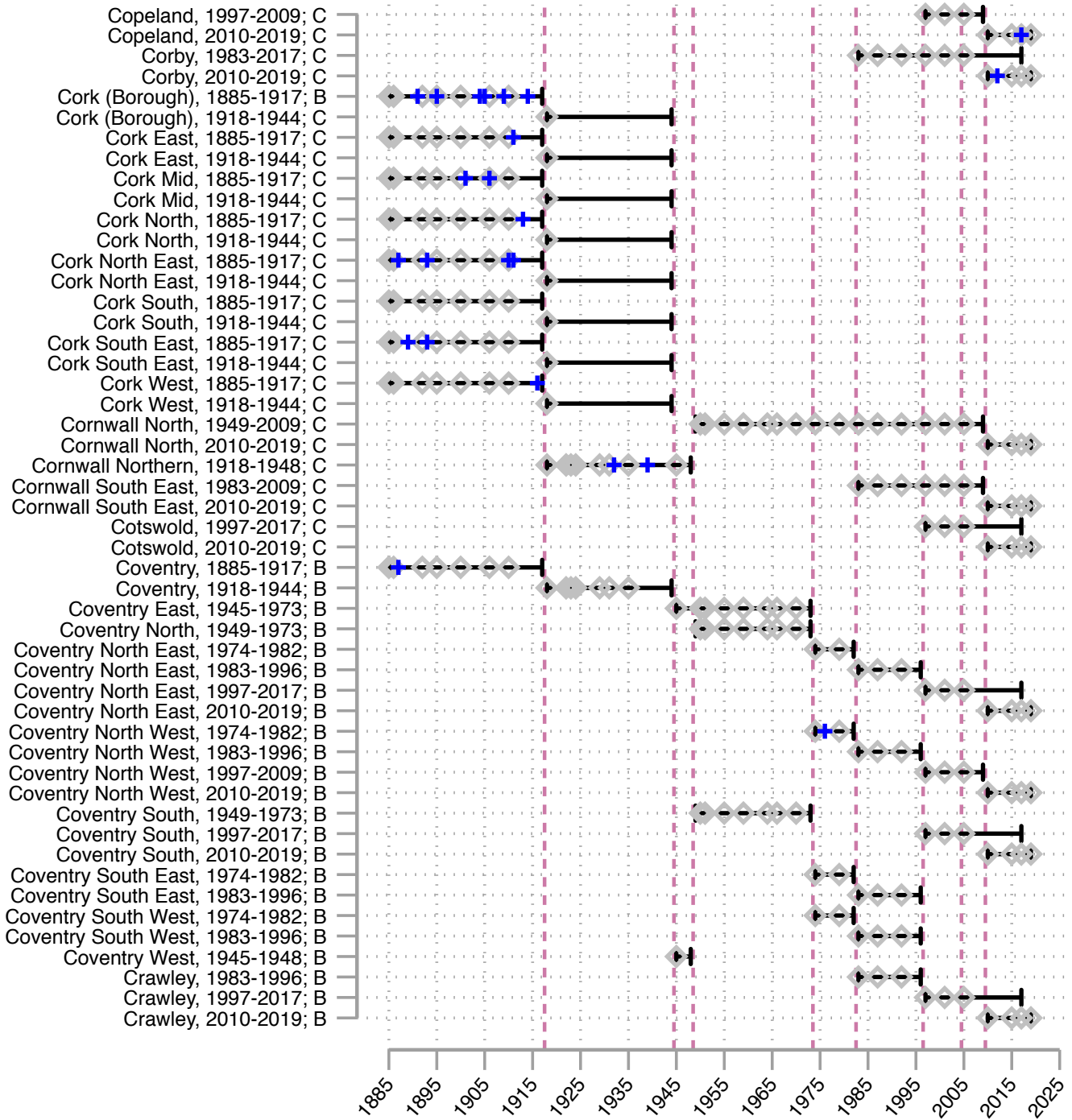


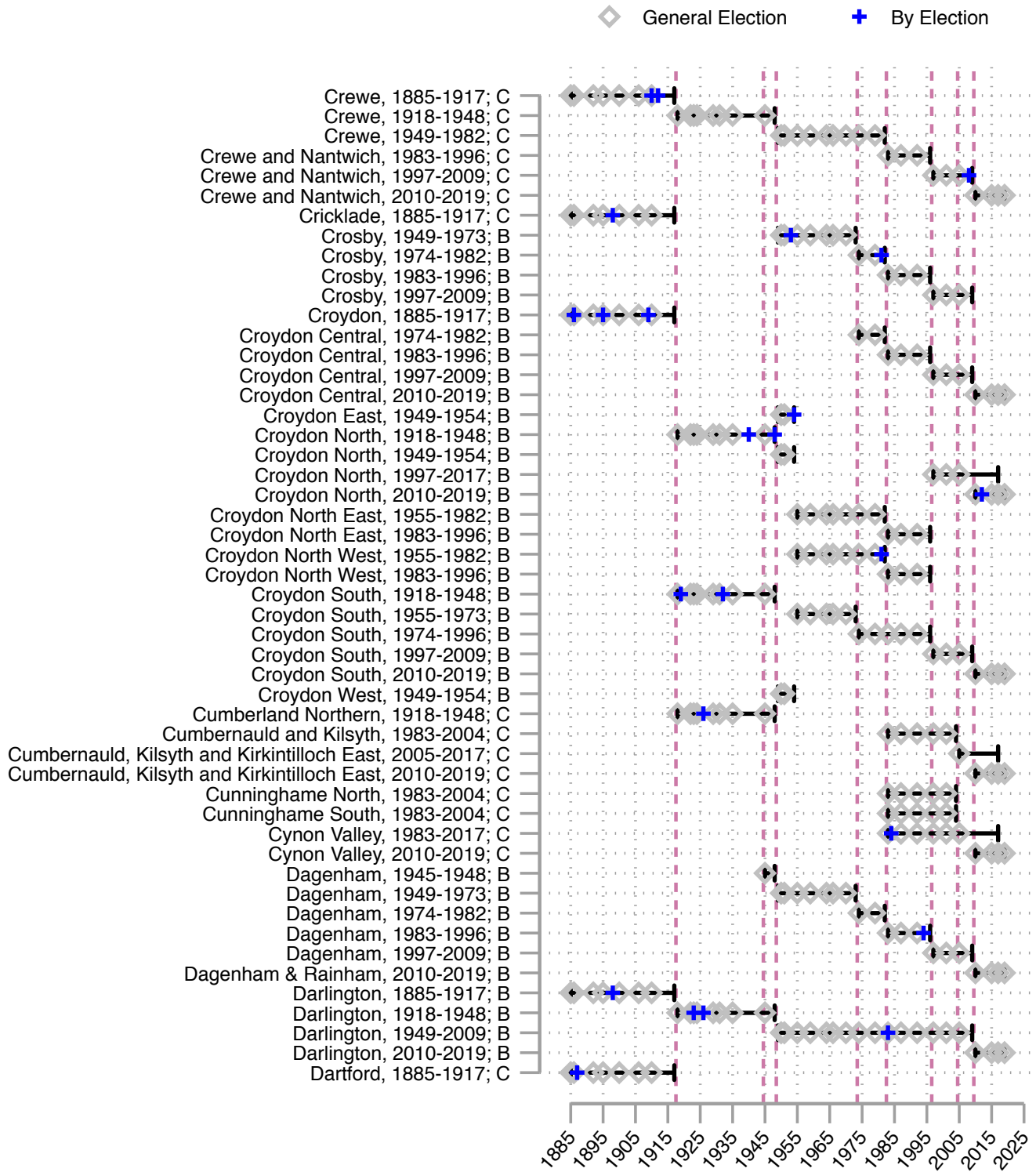


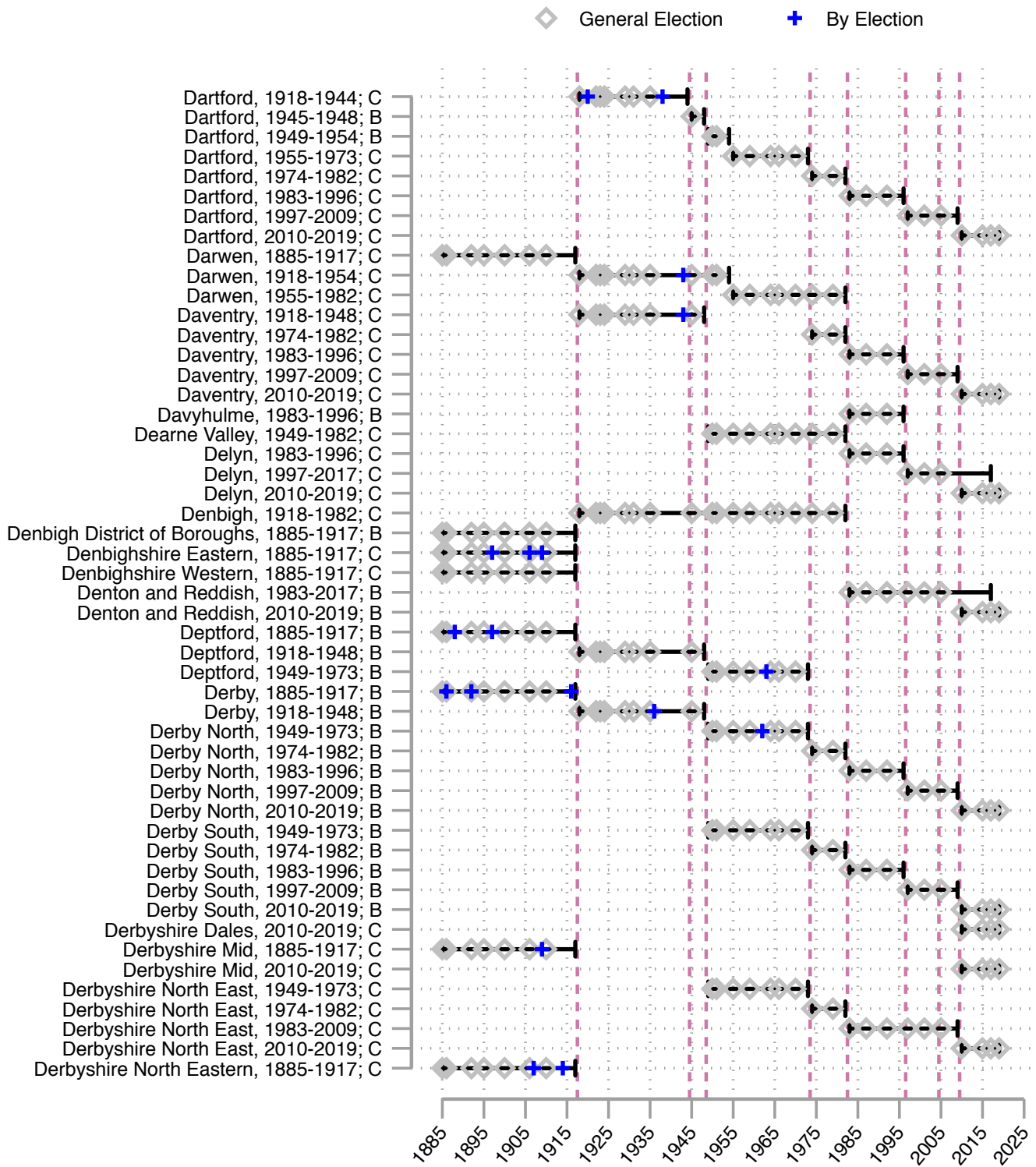
◇ General Election + By Election



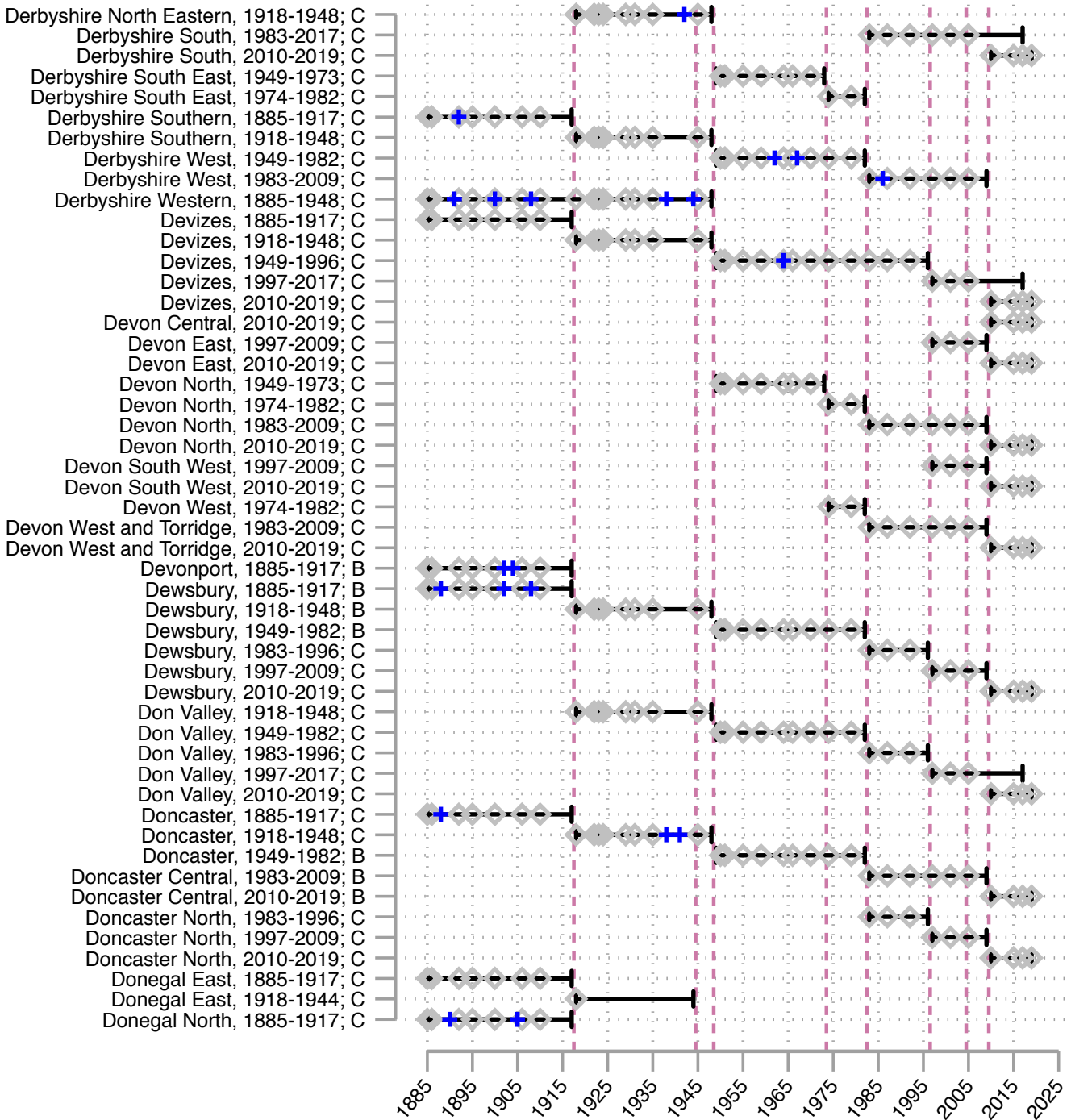
◇ General Election + By Election



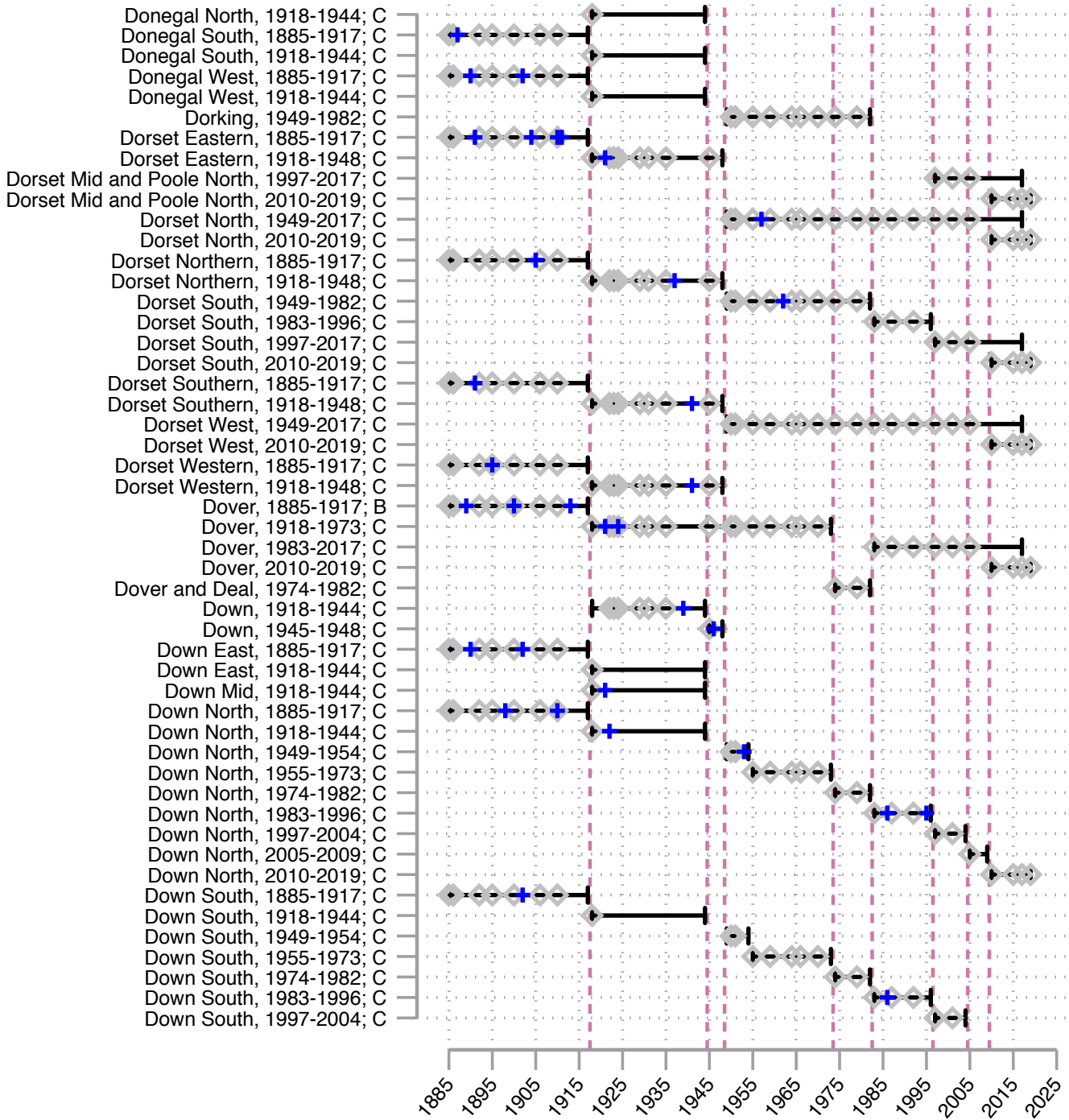




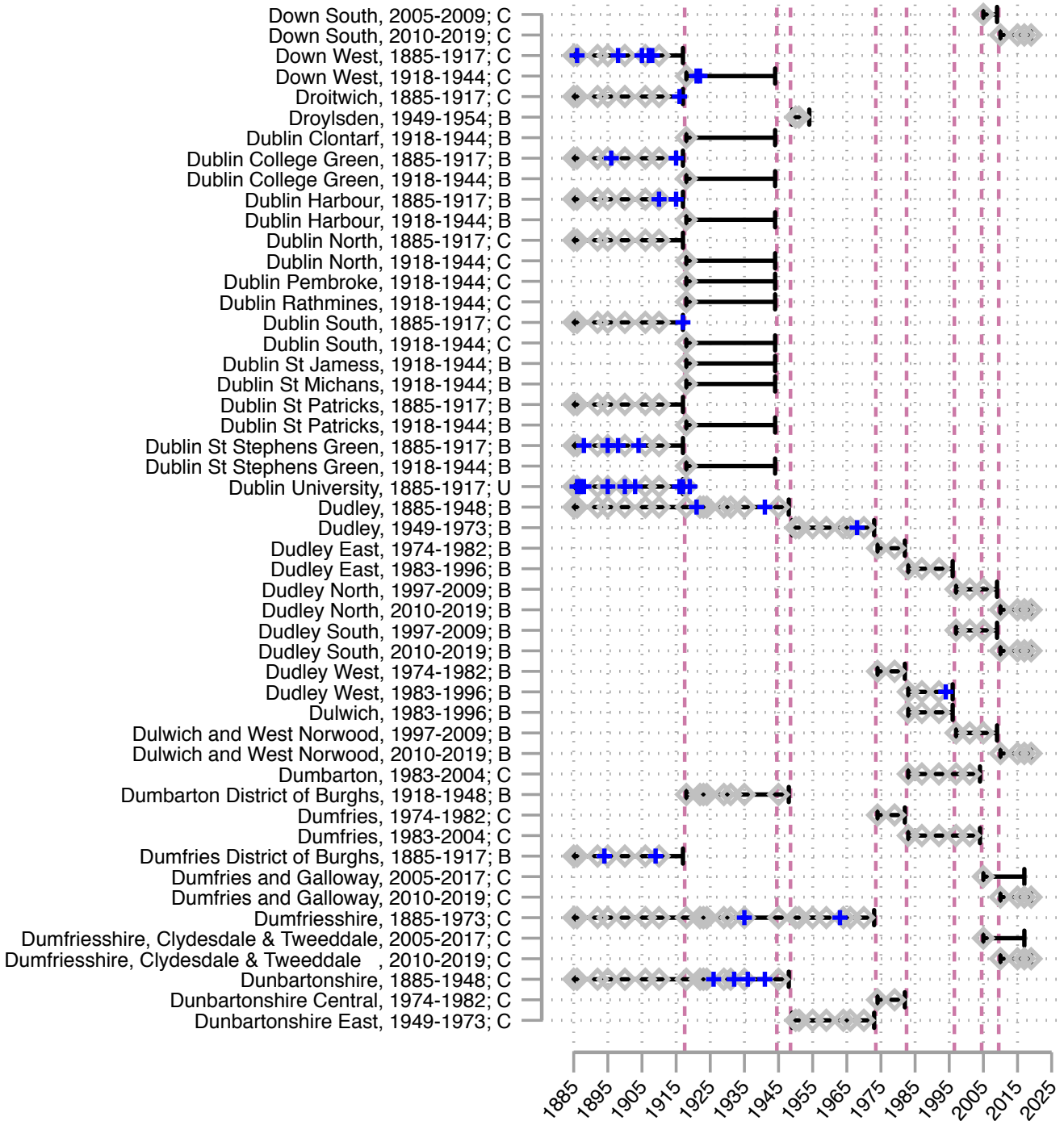
◇ General Election + By Election



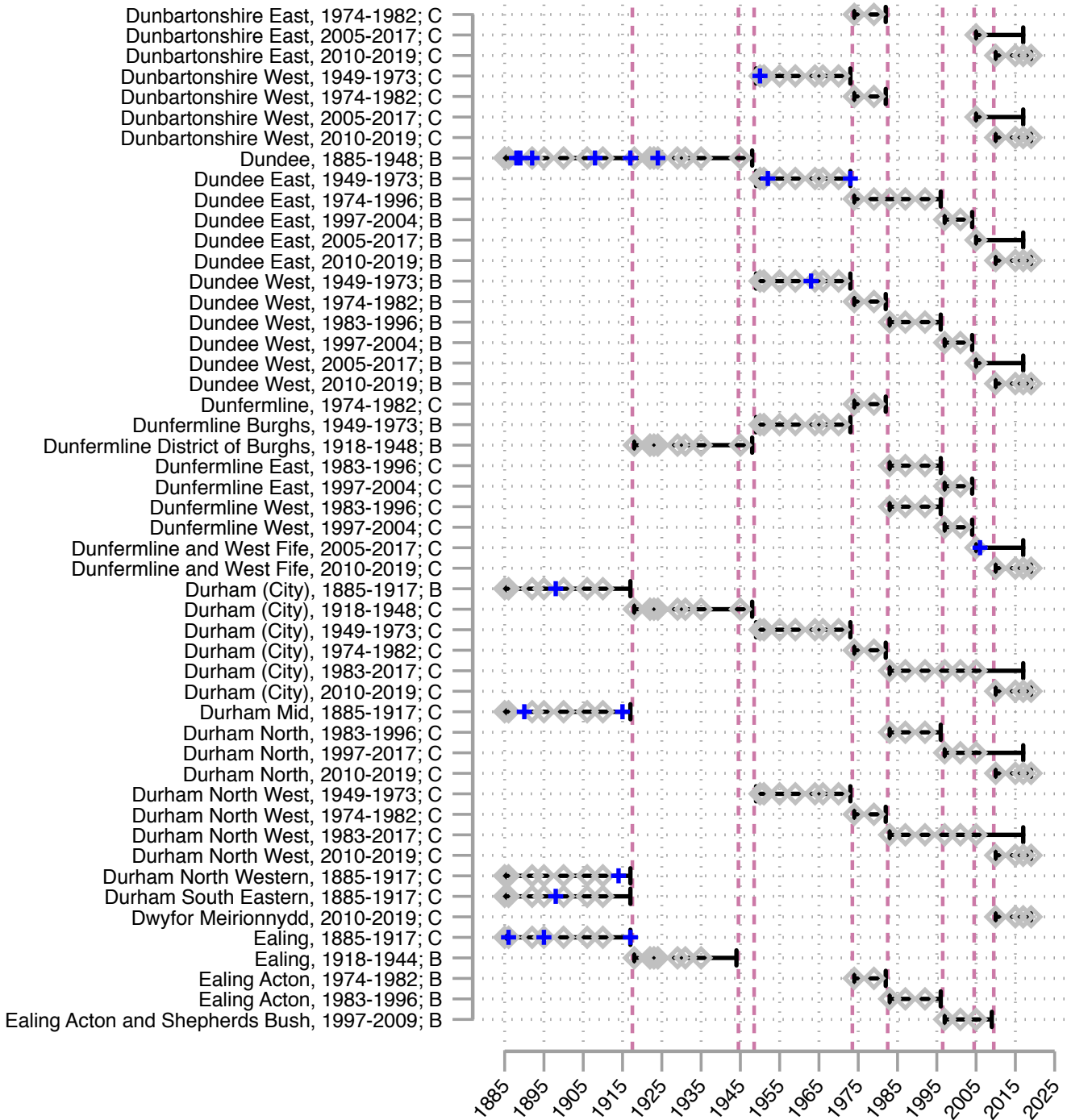
◇ General Election + By Election

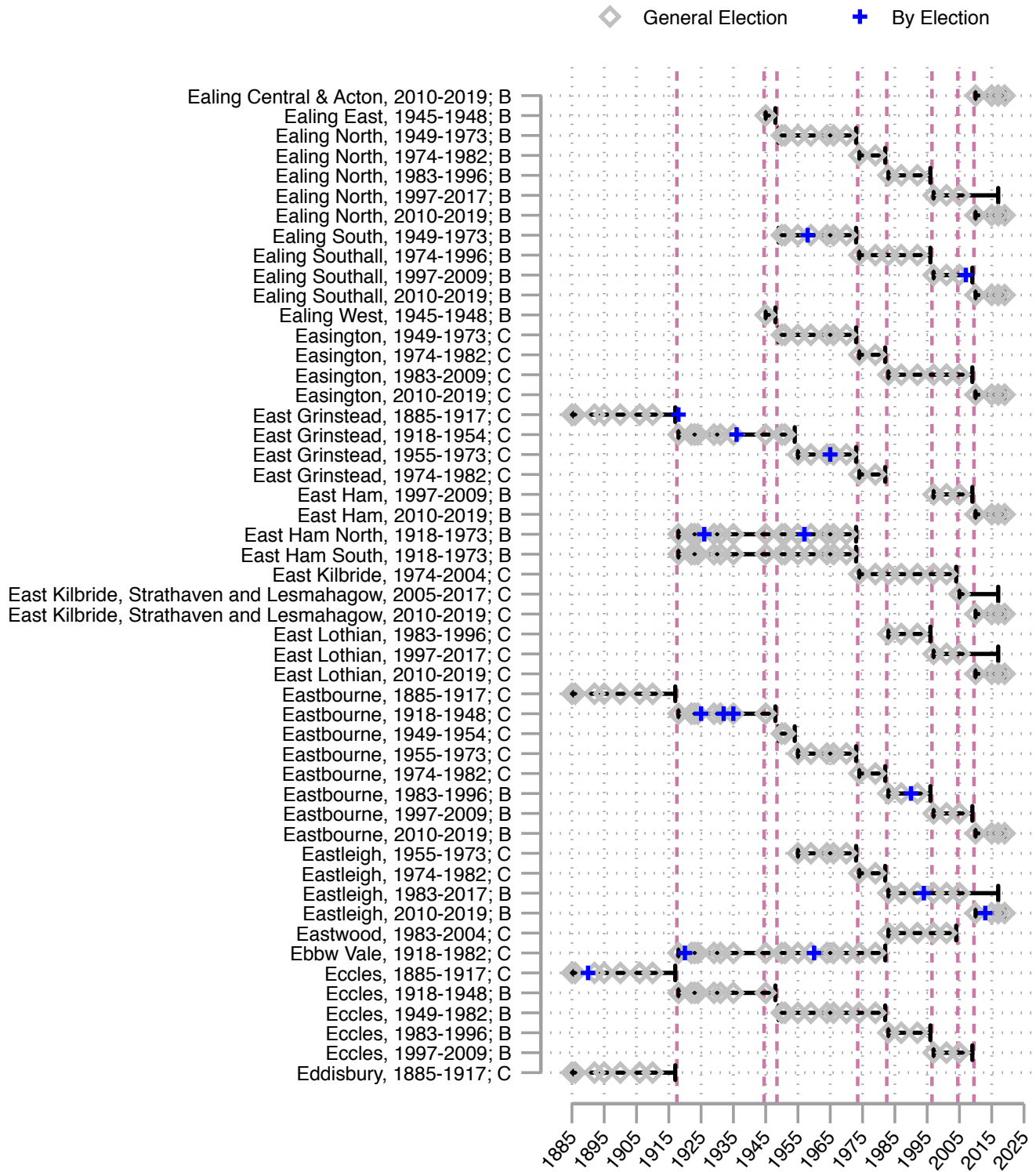


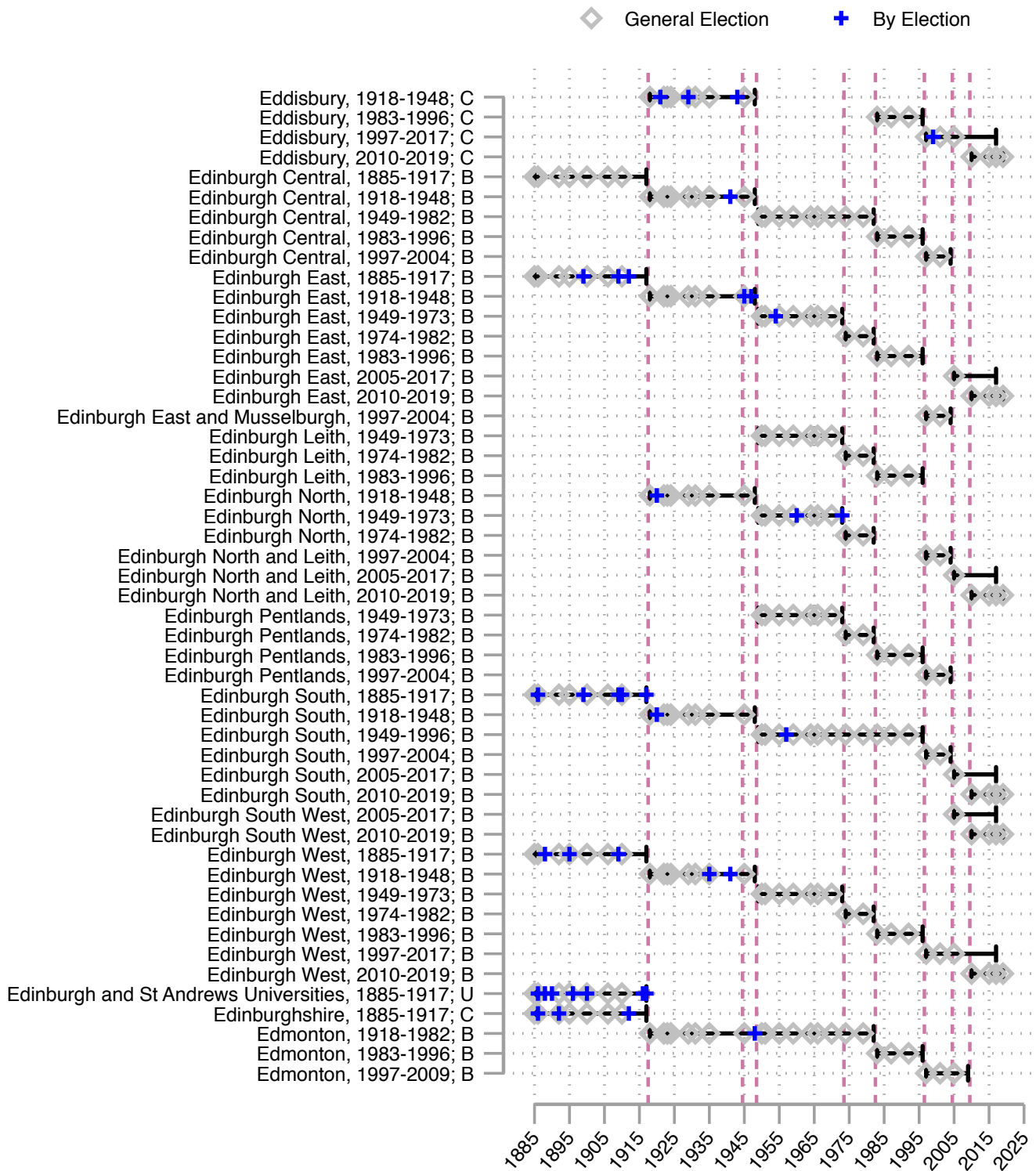
◇ General Election + By Election



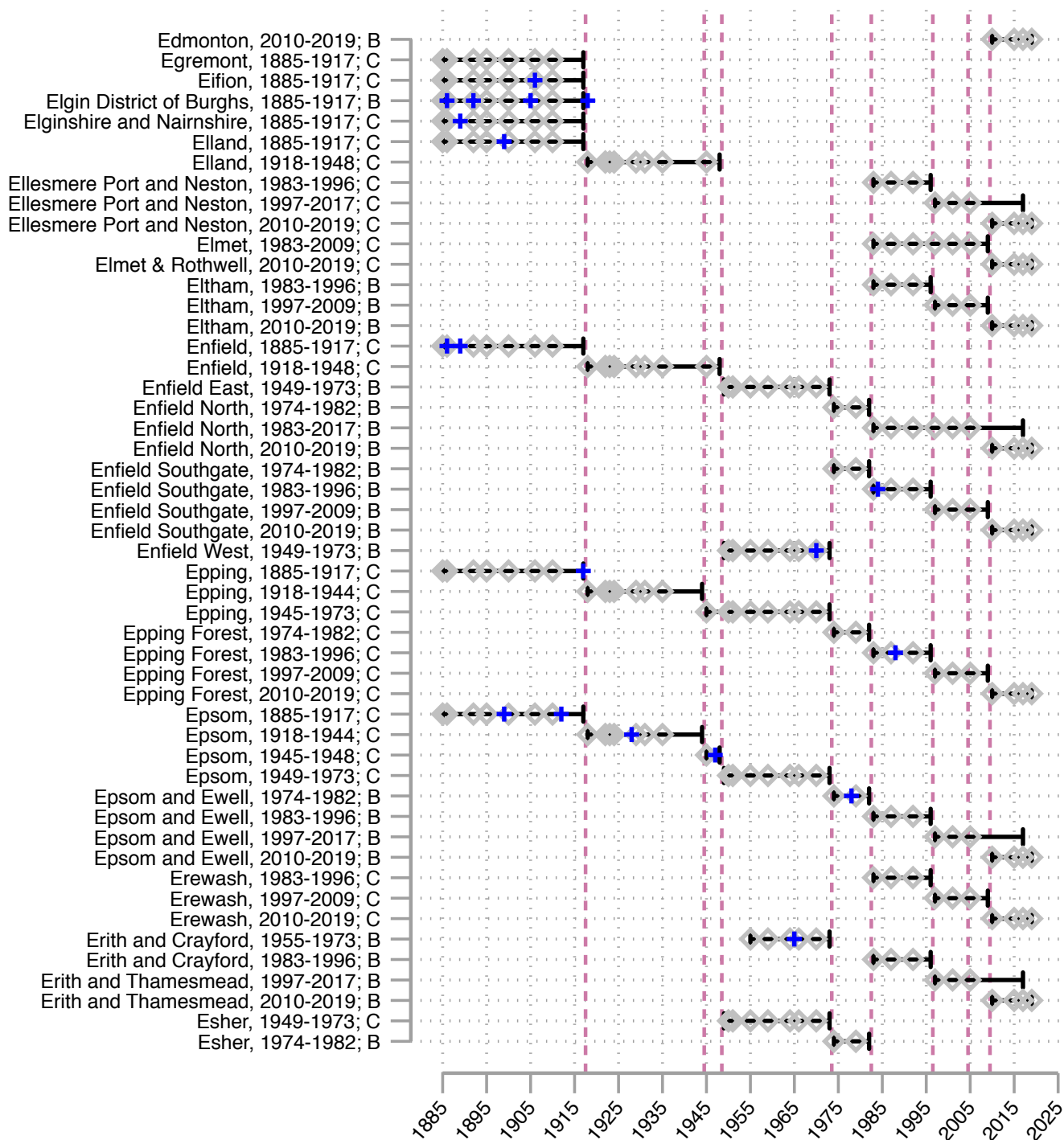
◇ General Election + By Election



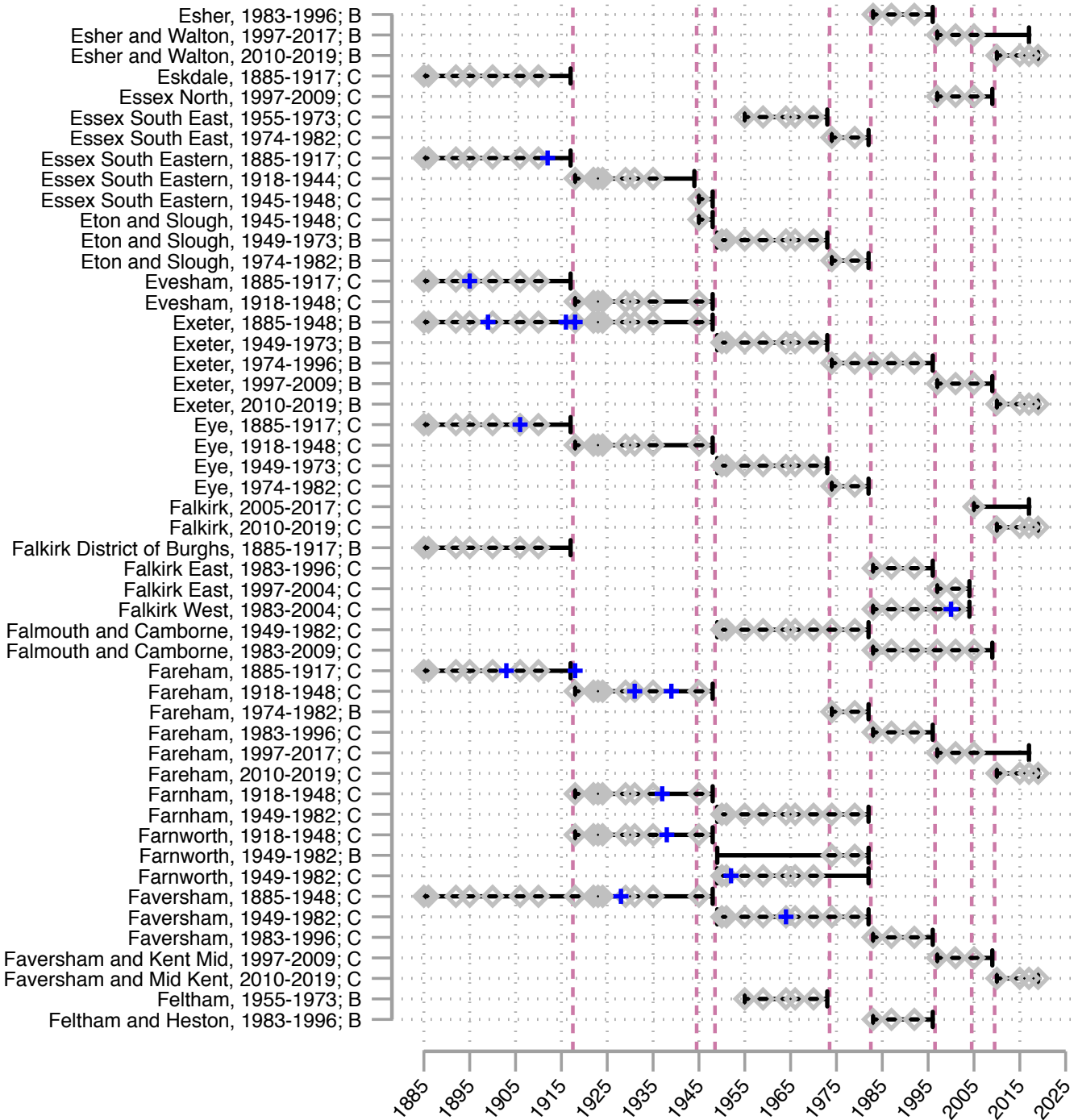




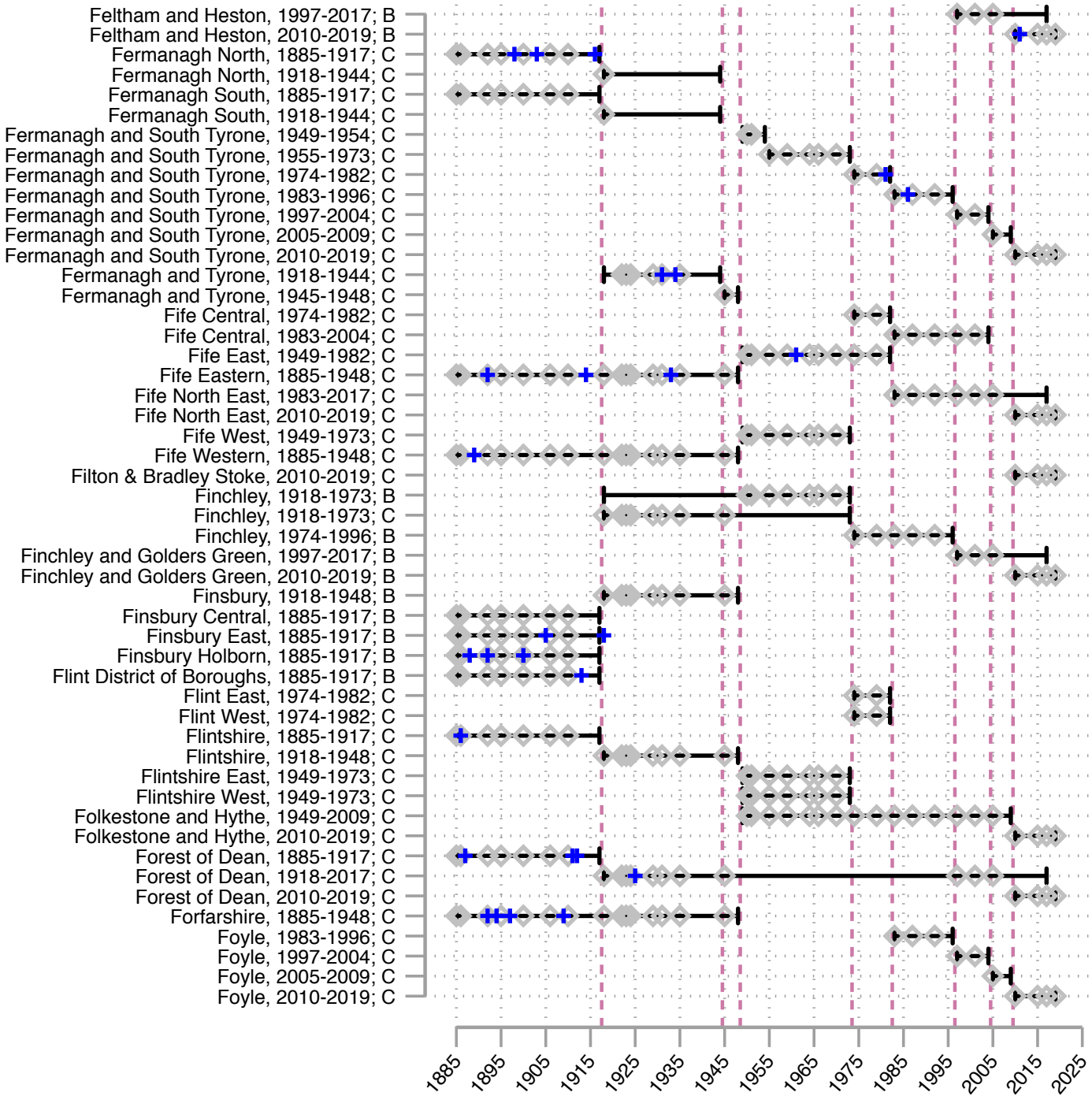
◇ General Election + By Election

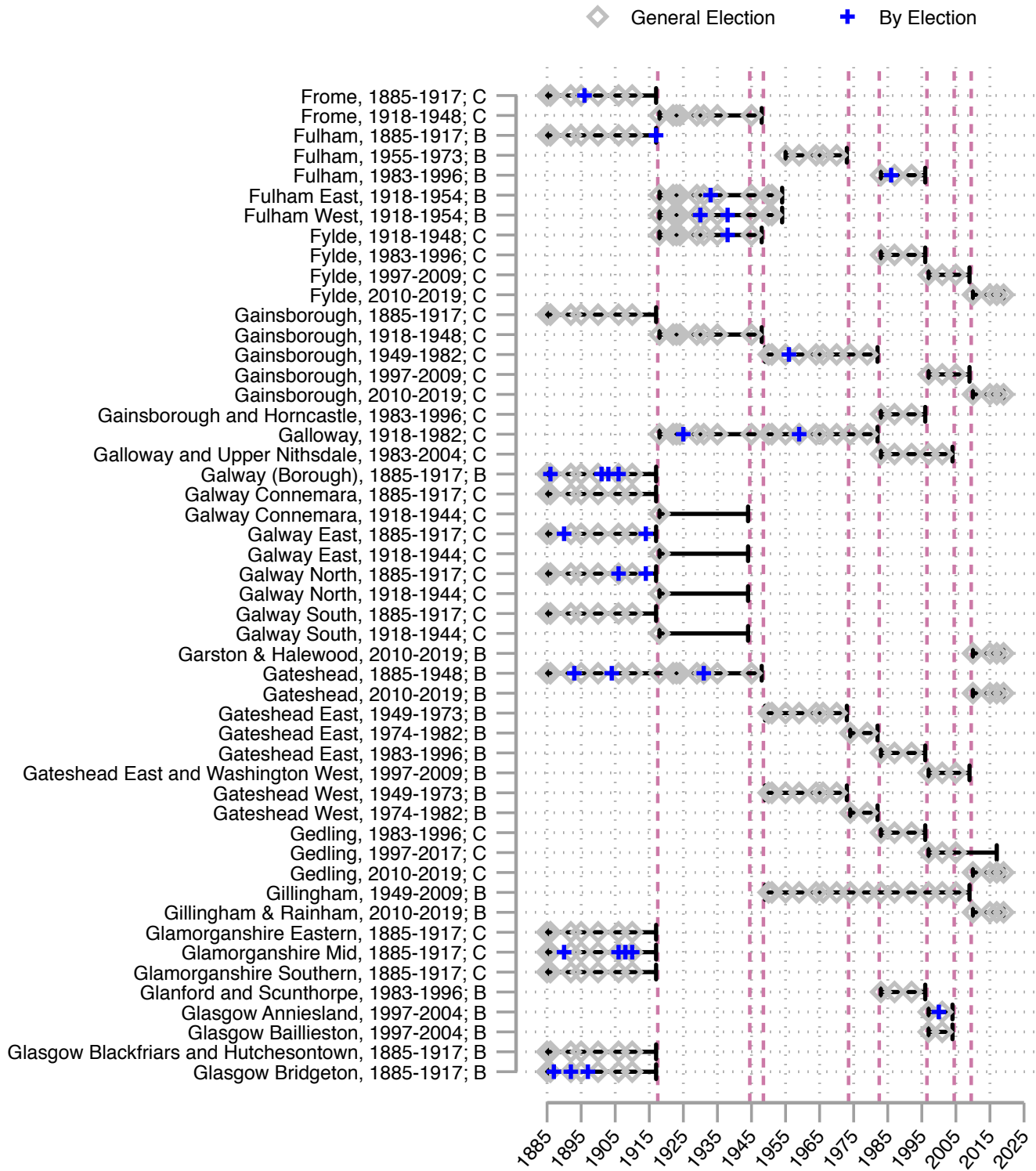


◇ General Election + By Election

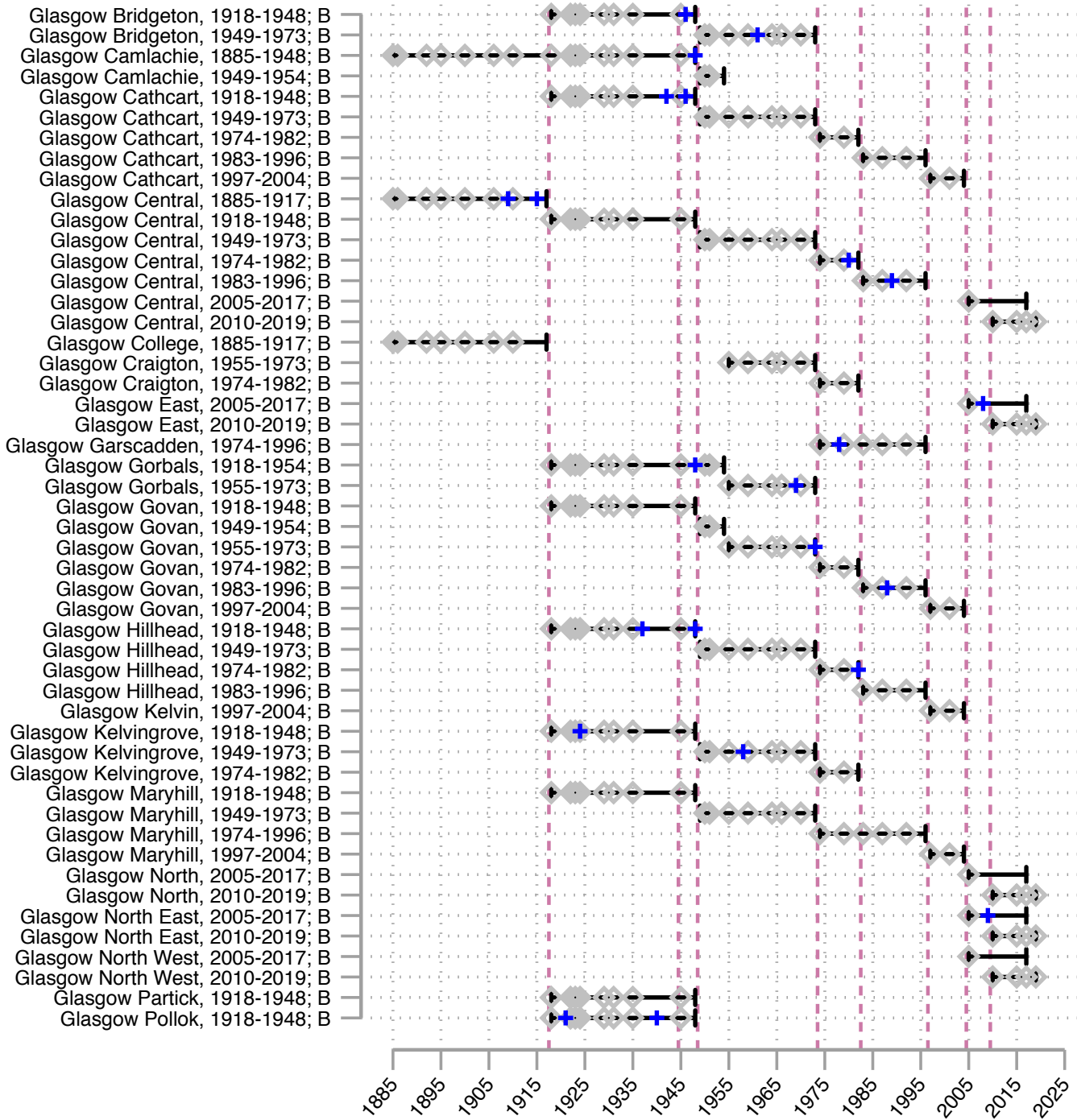


◇ General Election + By Election

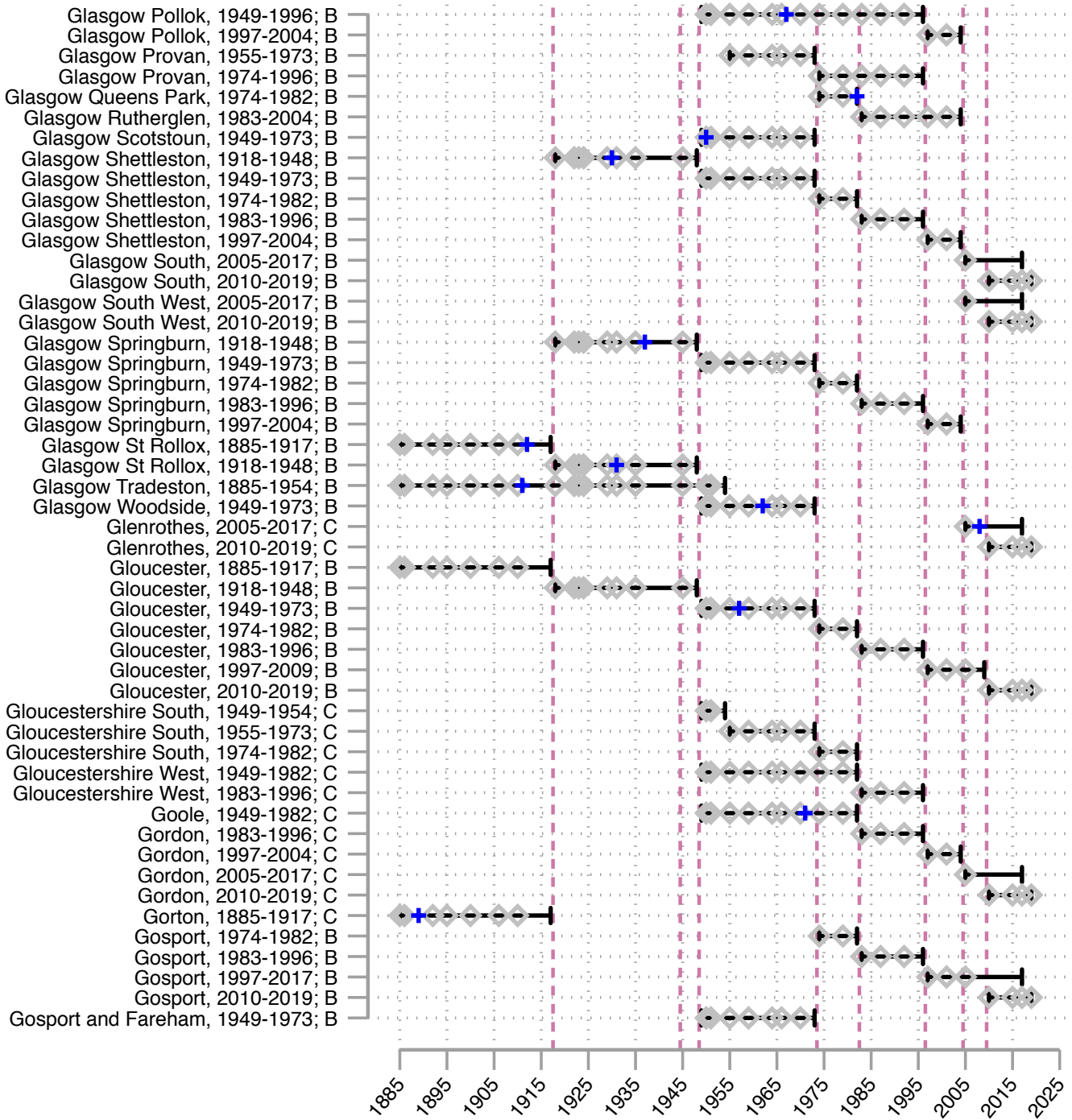


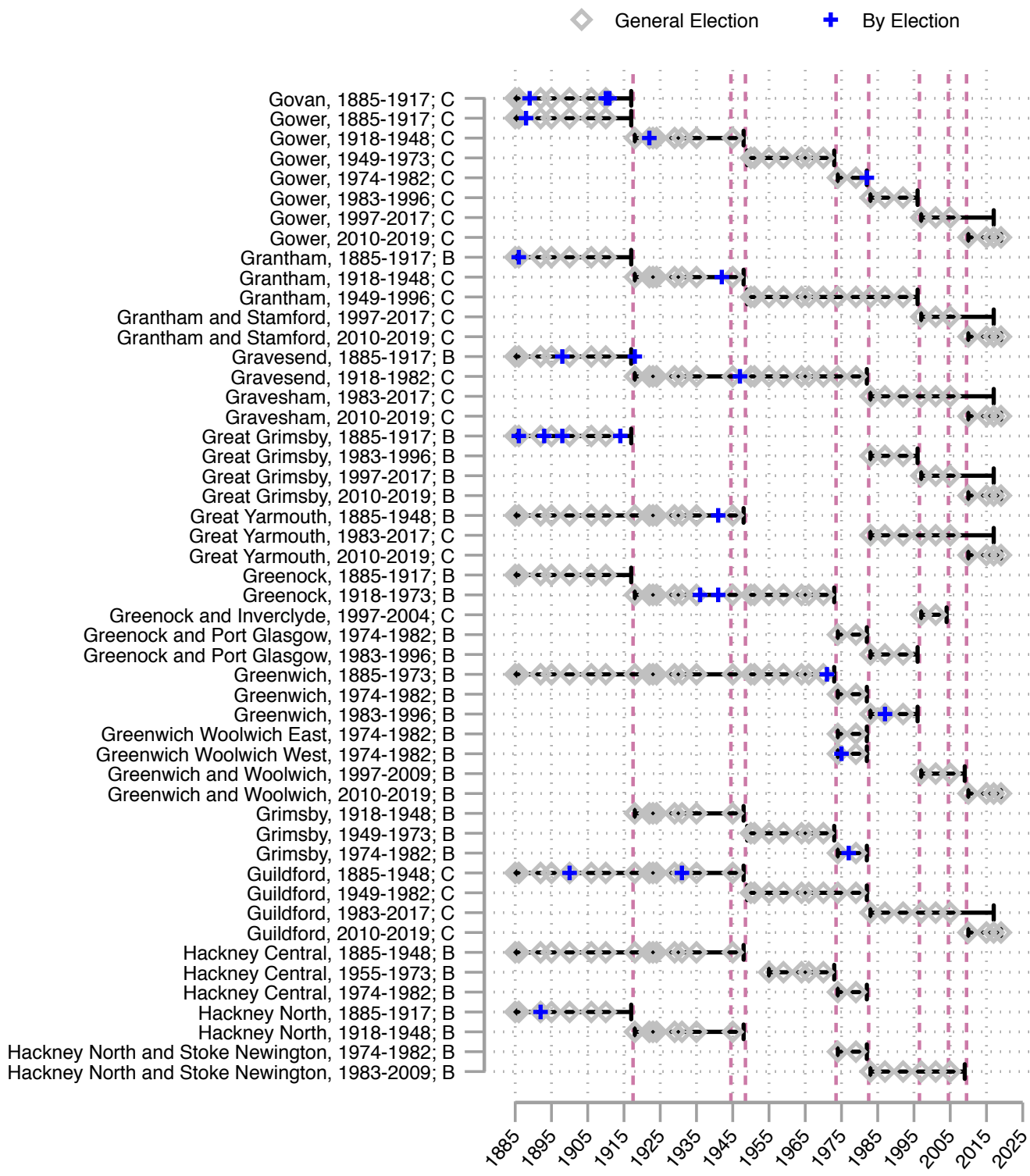


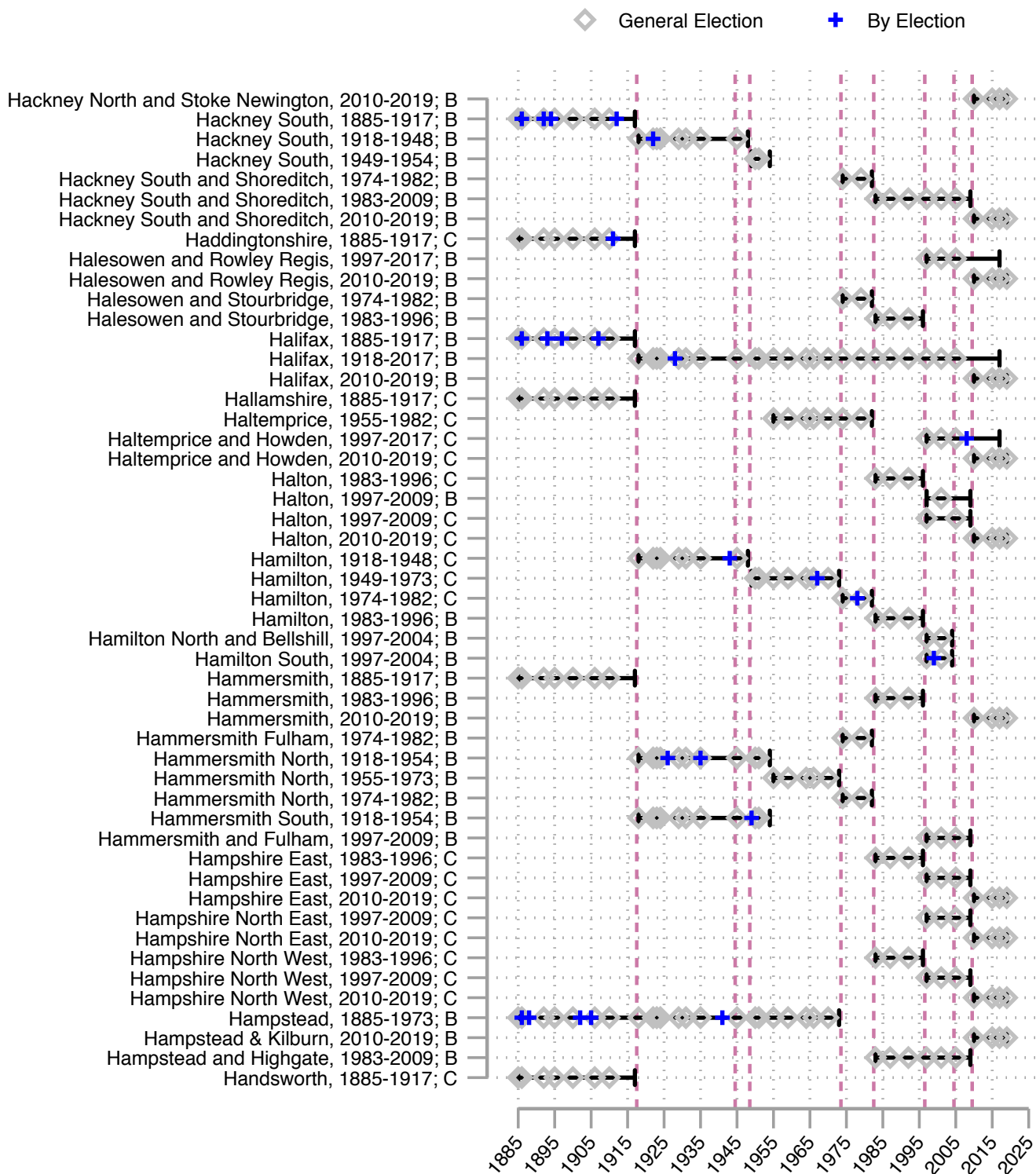
◇ General Election + By Election



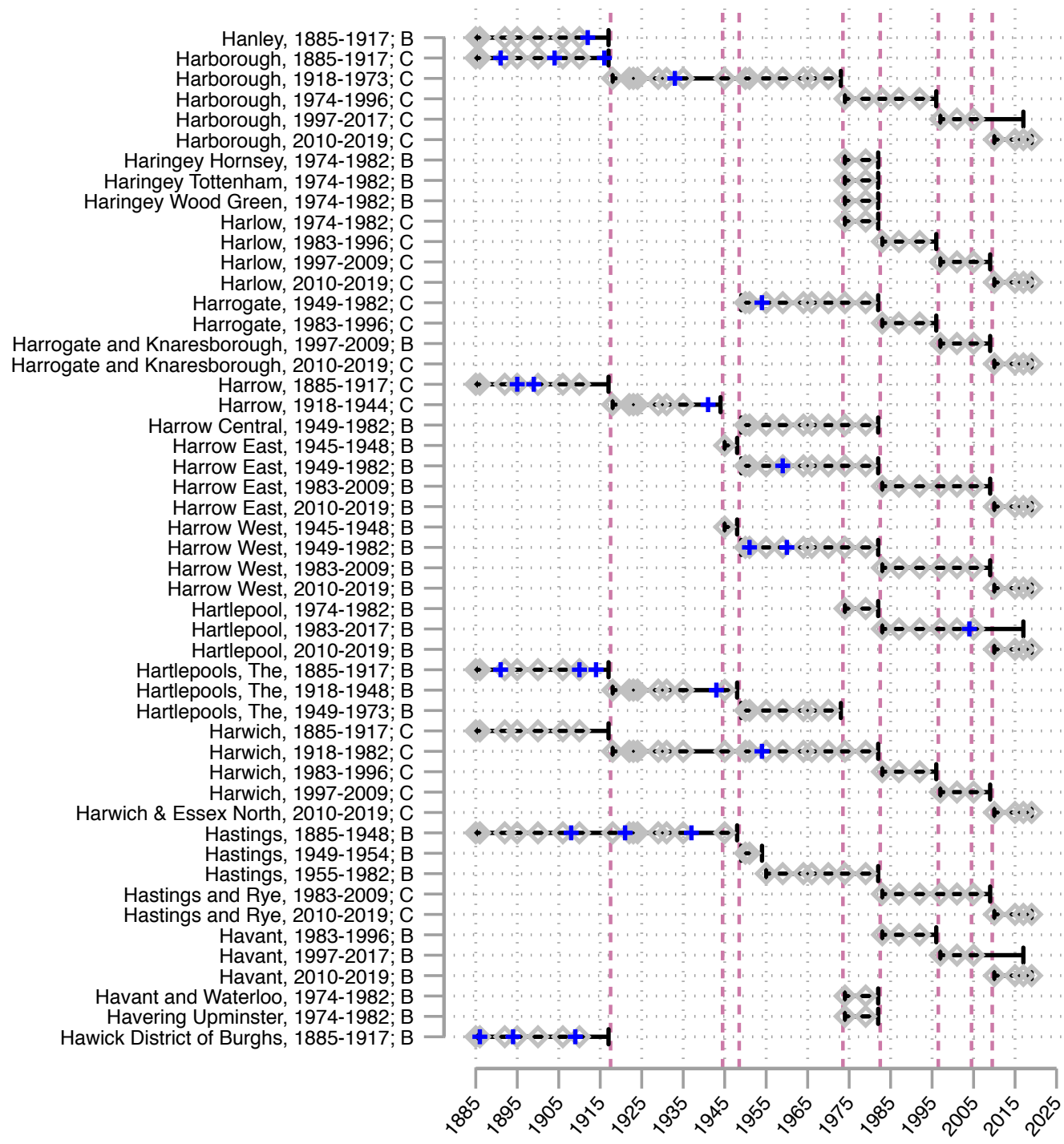
◇ General Election + By Election



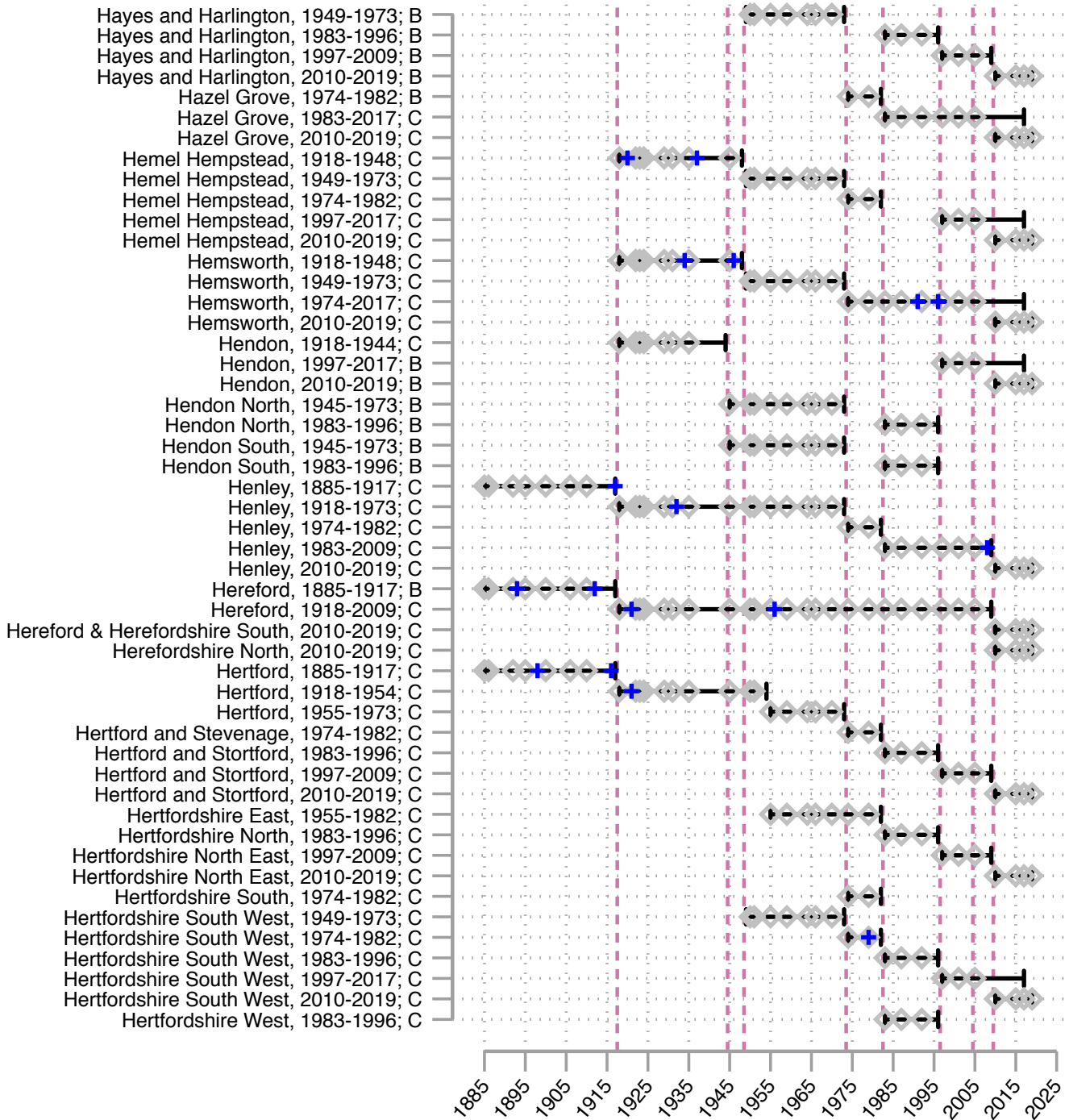




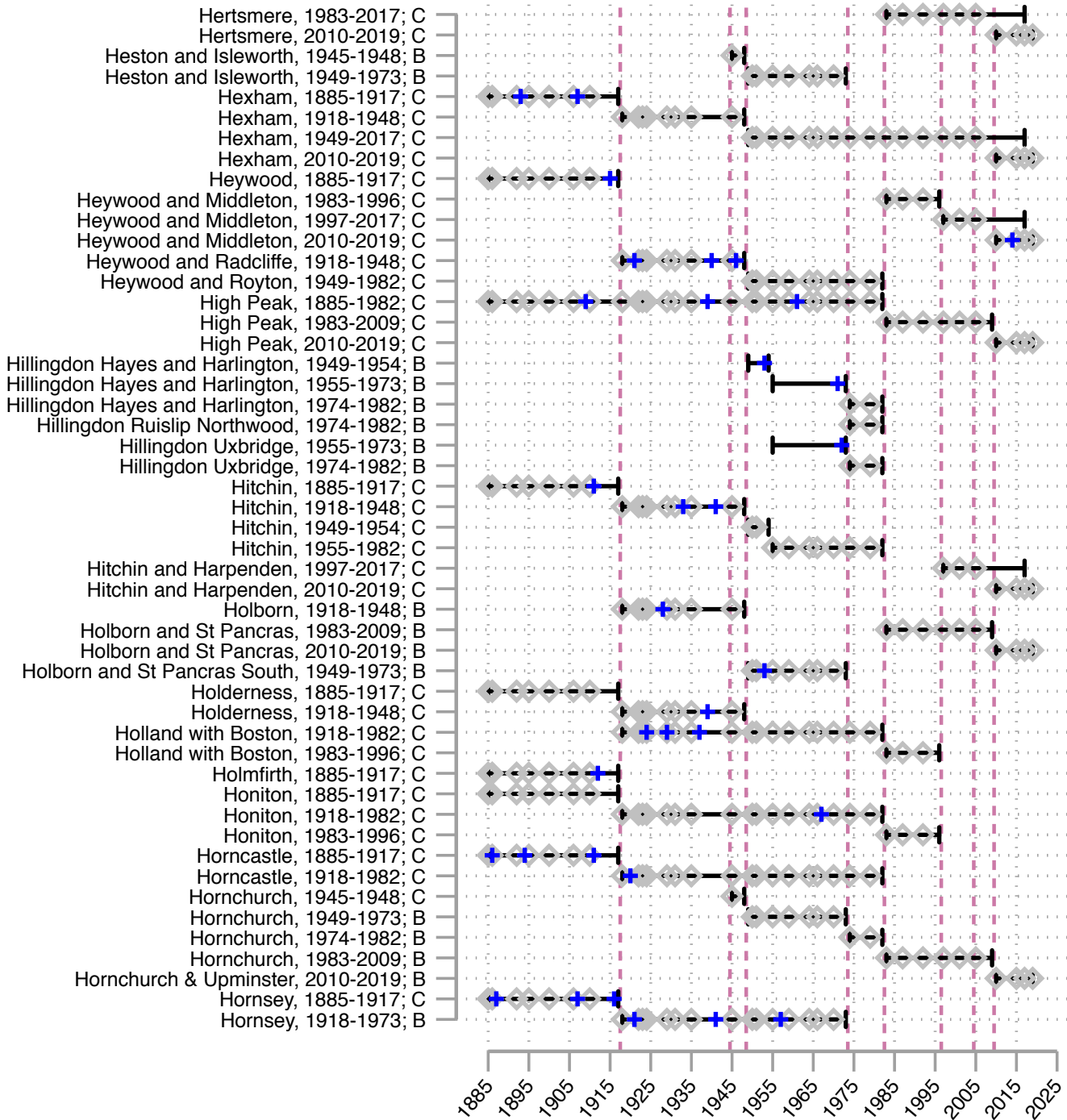
◇ General Election + By Election



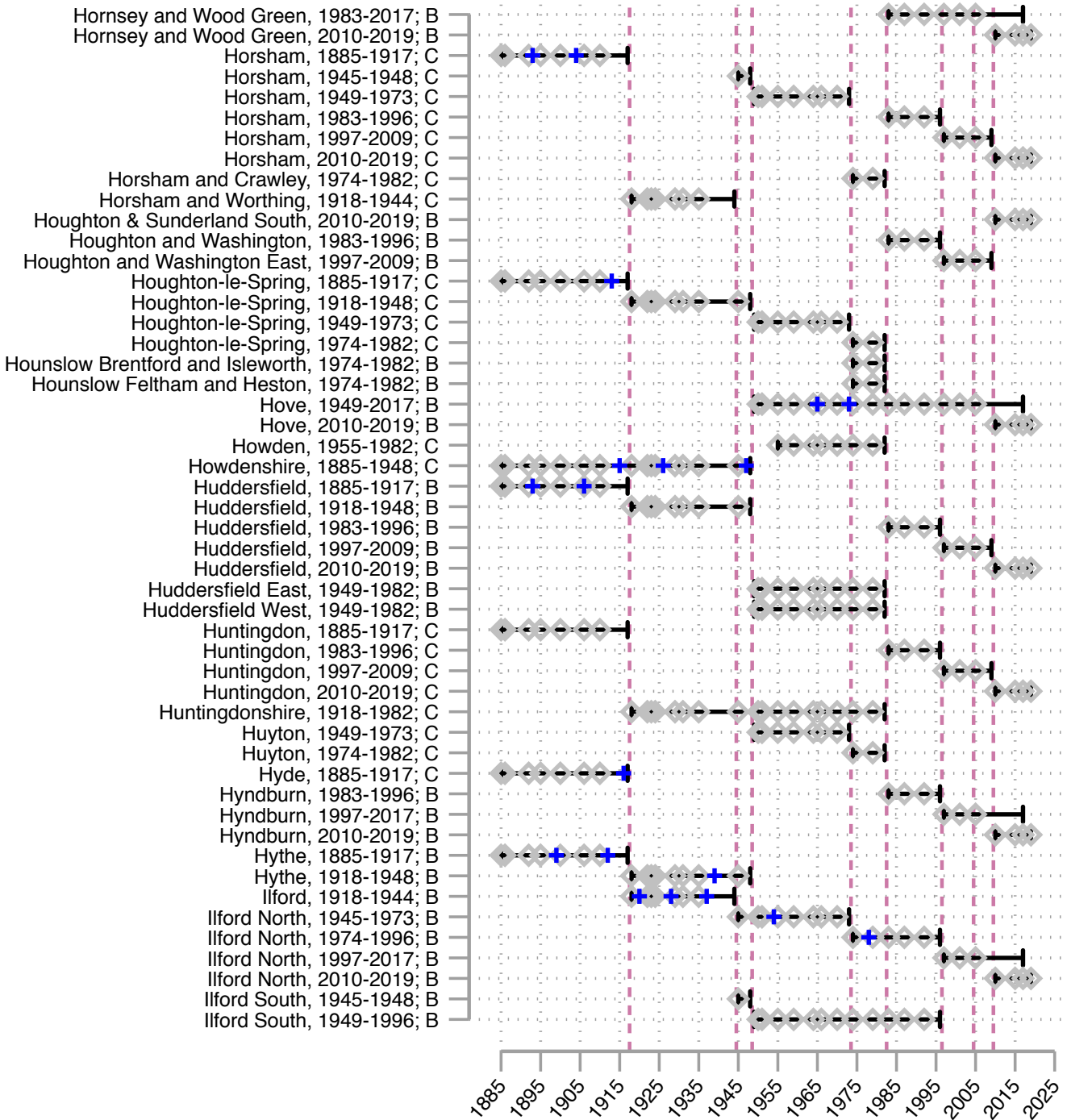
◇ General Election + By Election

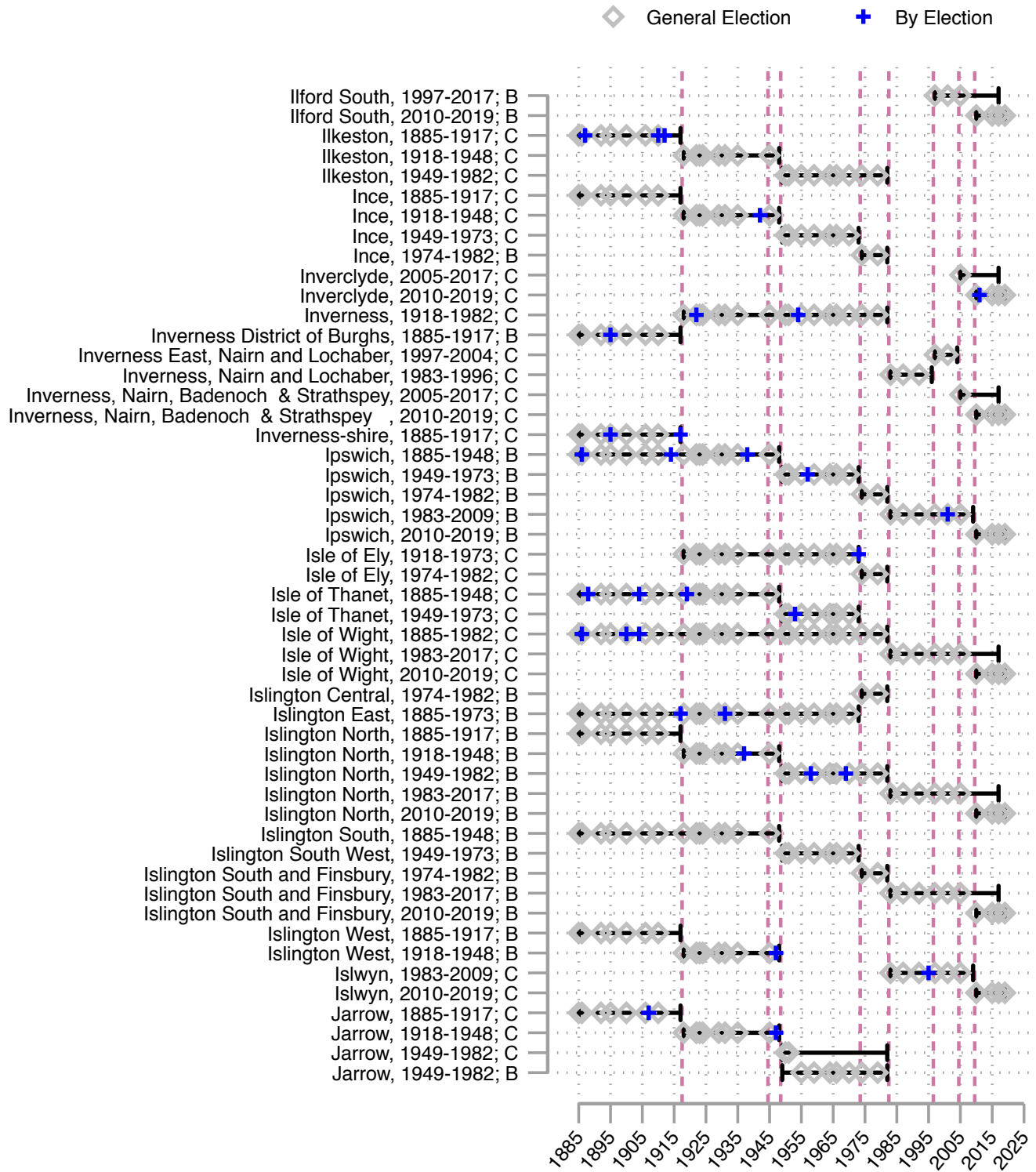


◇ General Election + By Election

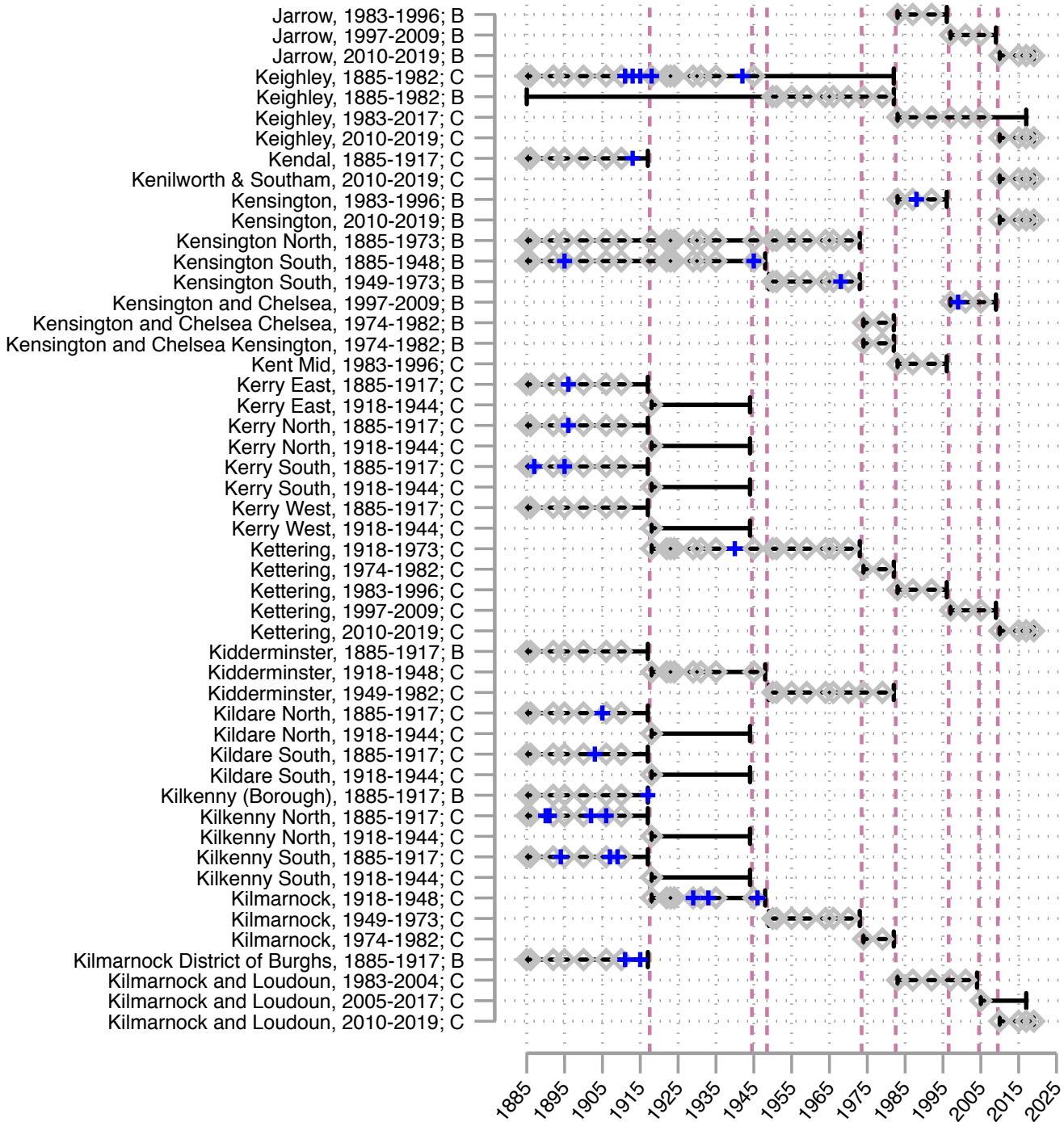


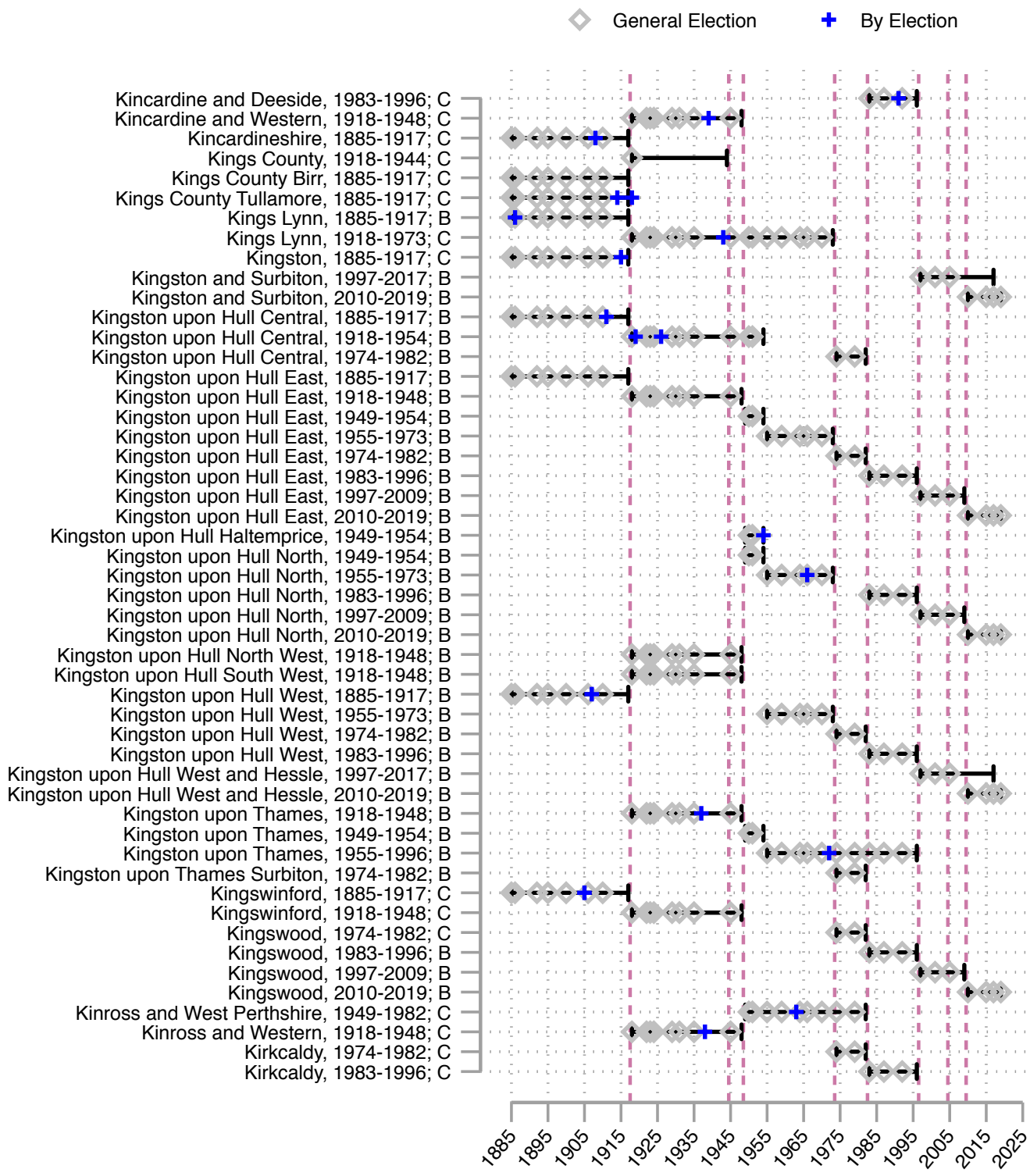
◇ General Election + By Election



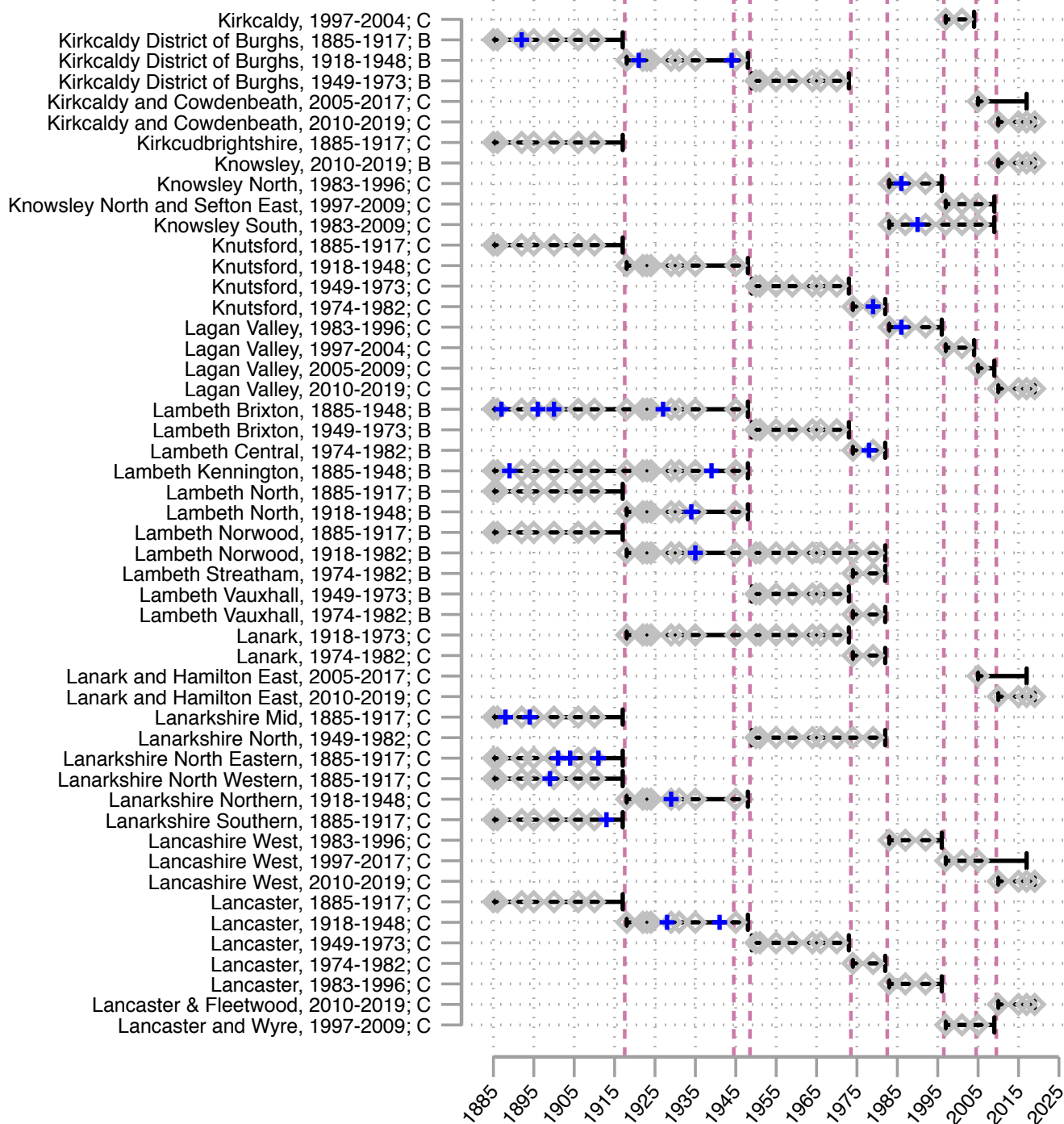


◇ General Election + By Election

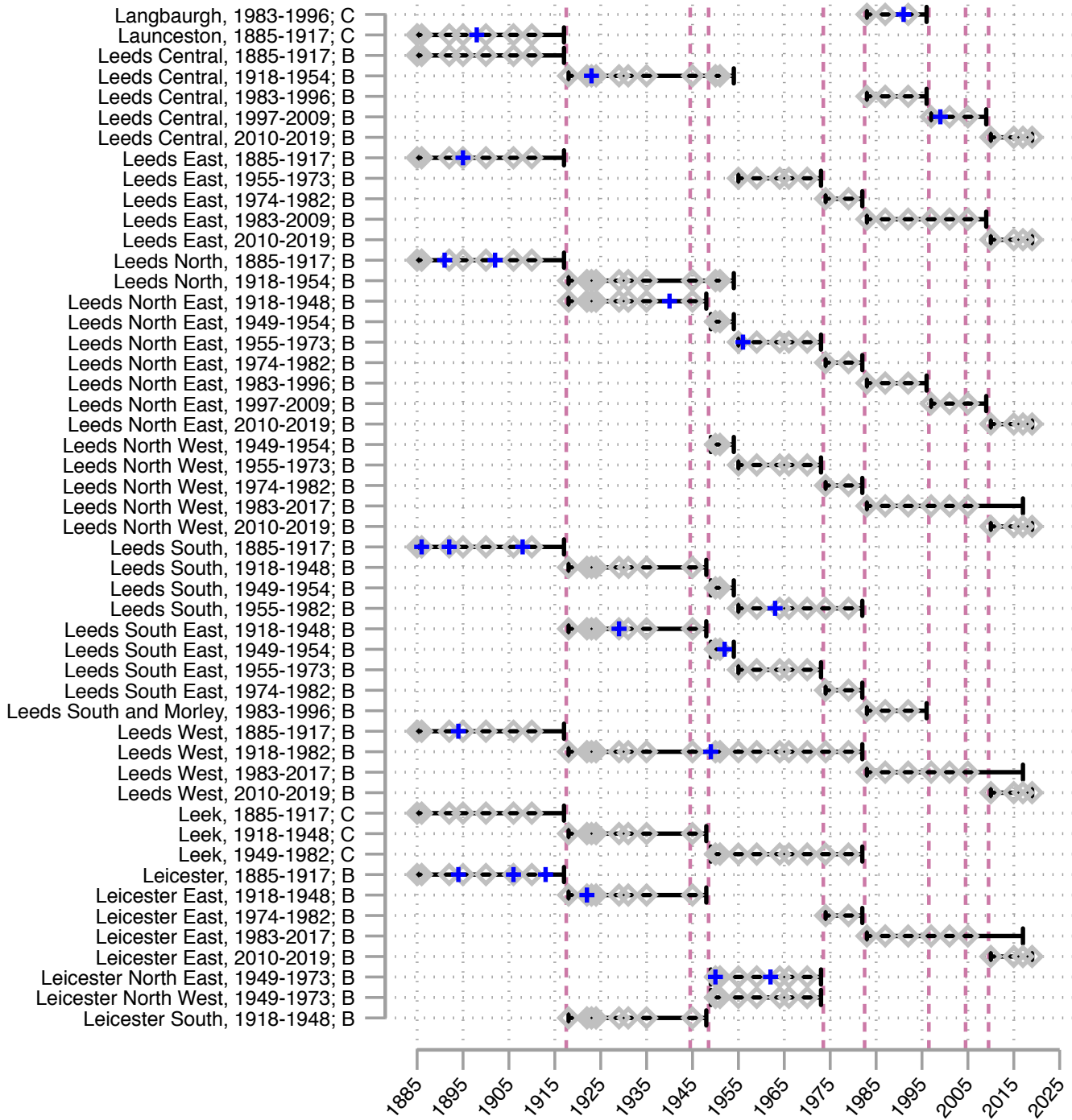




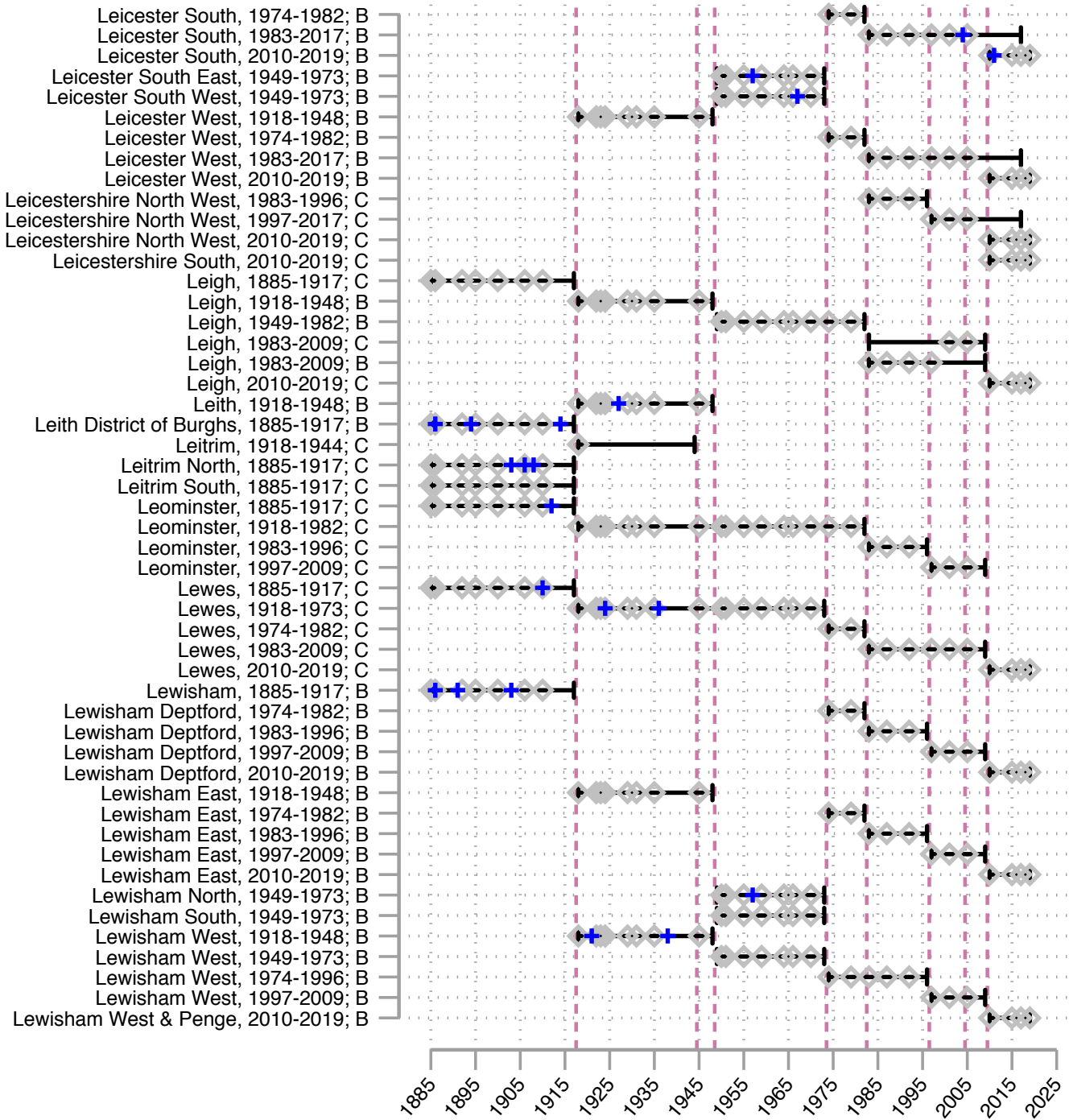
◇ General Election + By Election



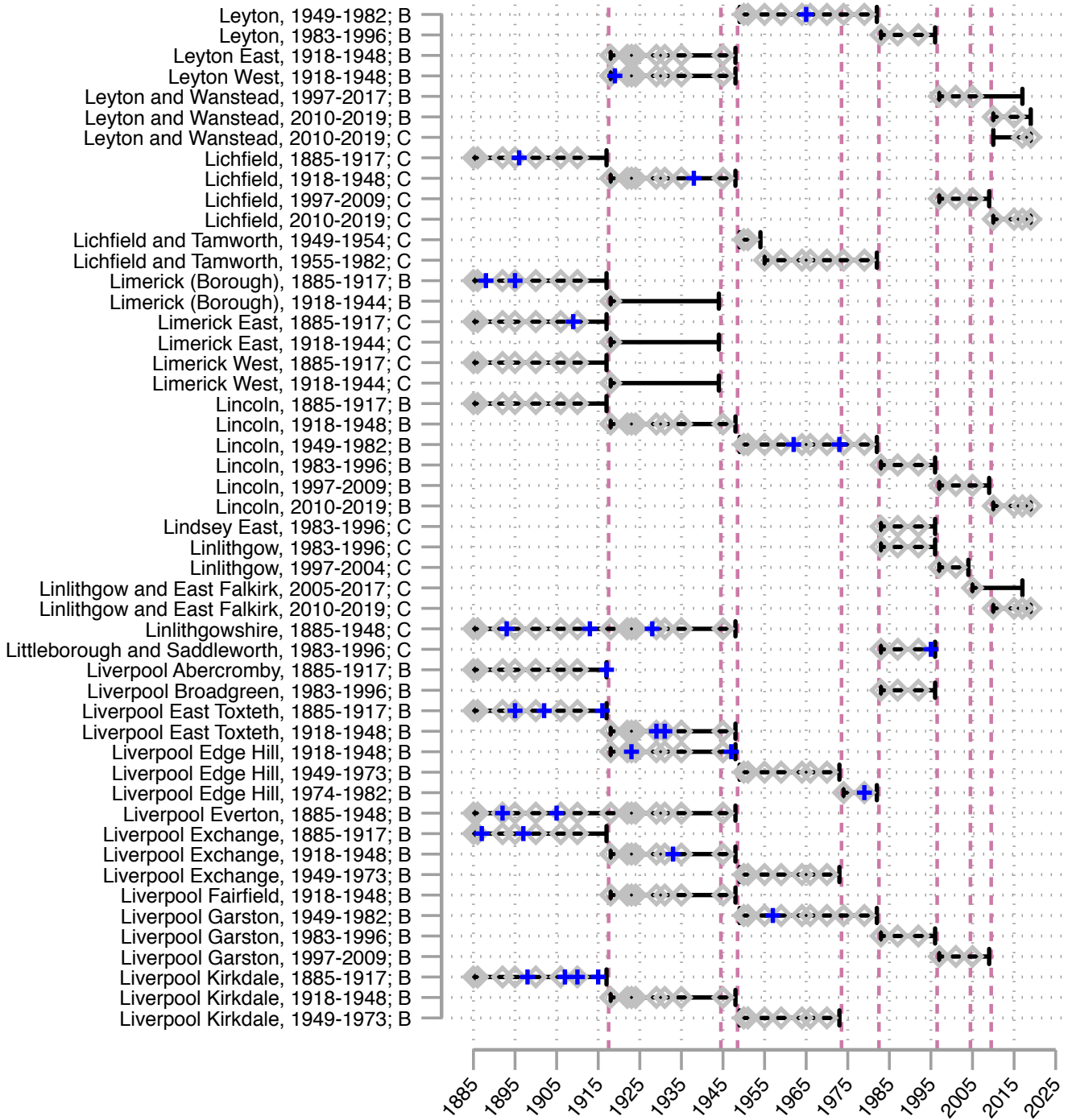
◇ General Election + By Election



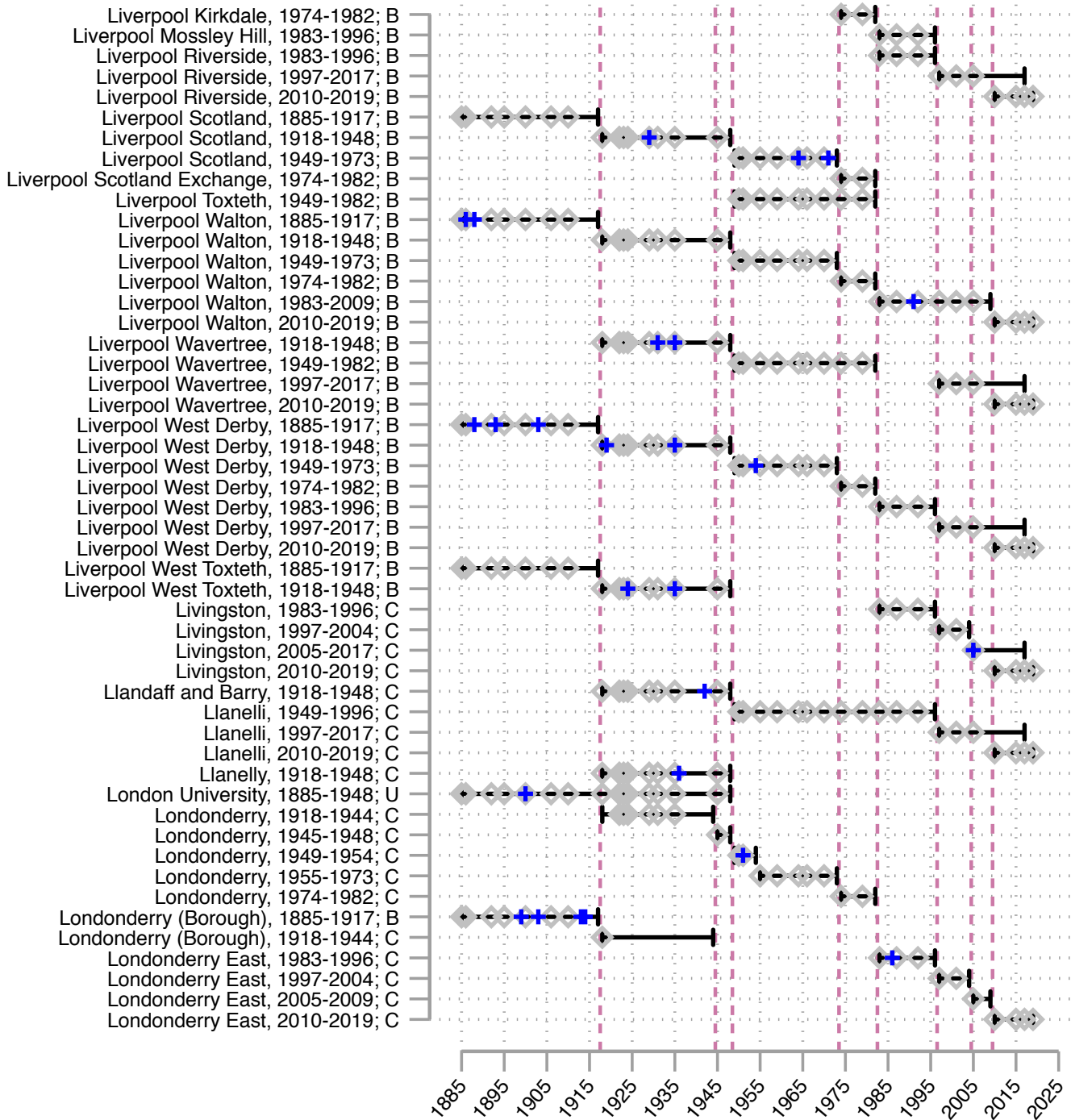
◇ General Election + By Election



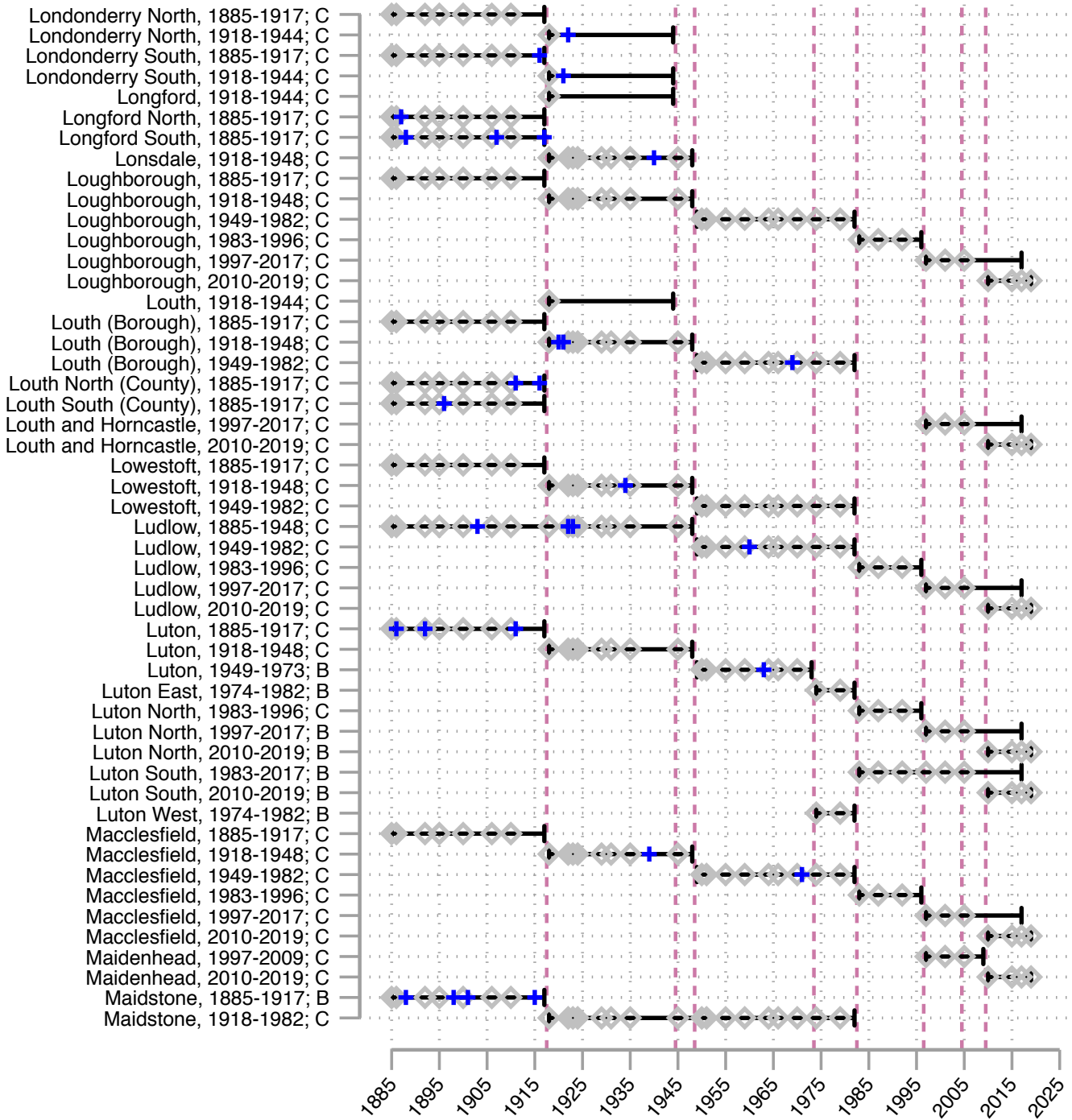
◇ General Election + By Election



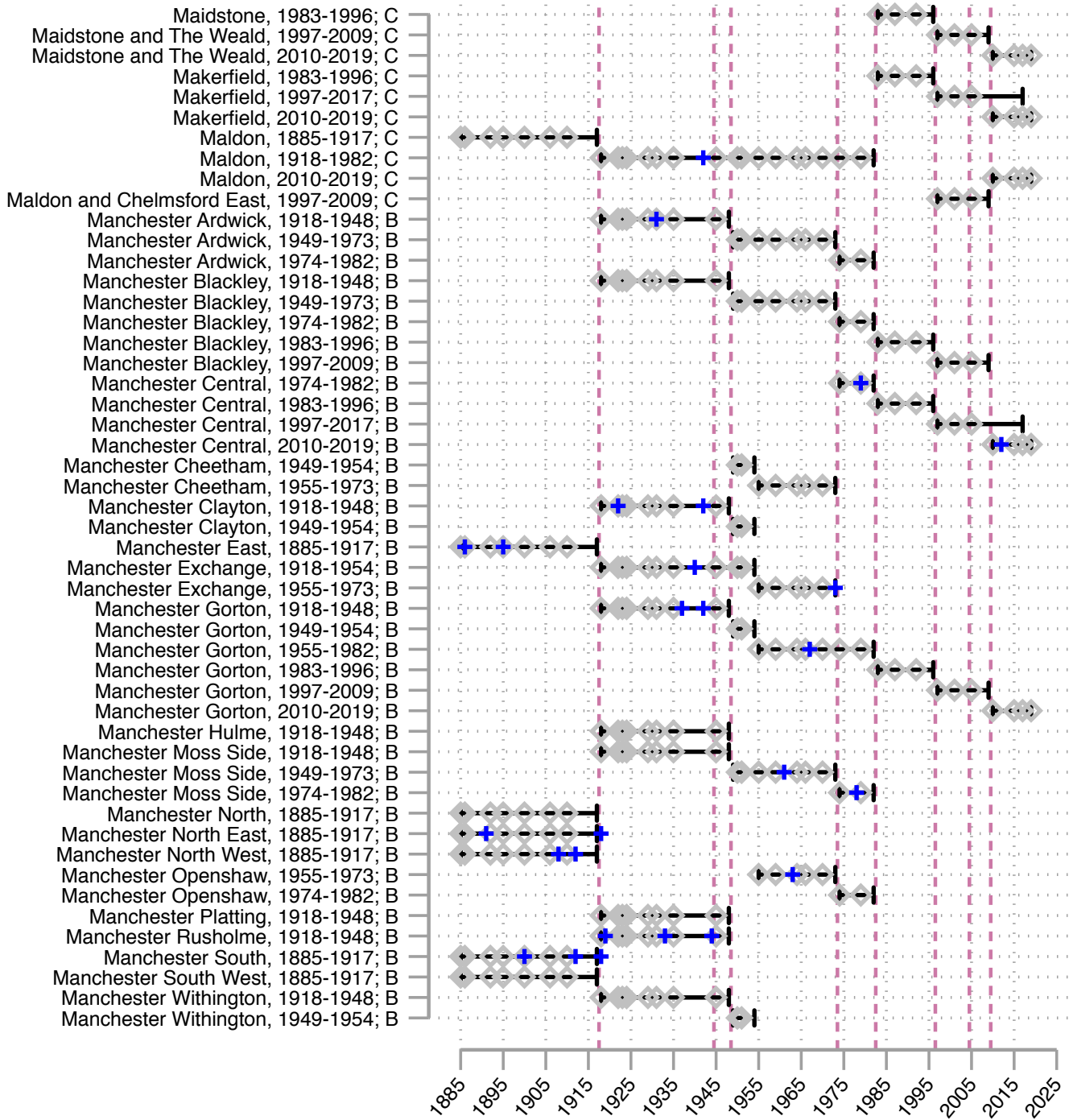
◇ General Election + By Election



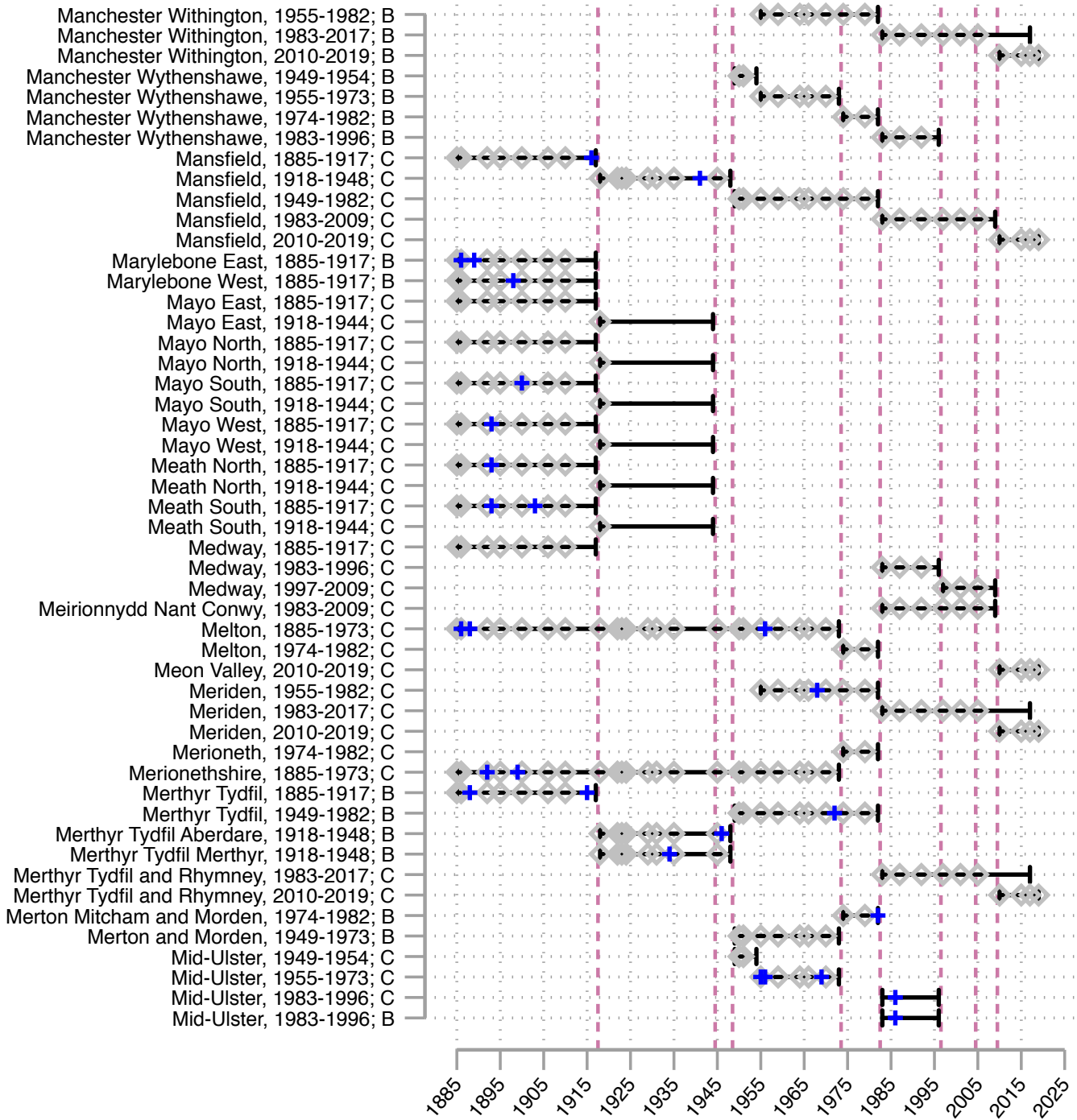
◇ General Election + By Election



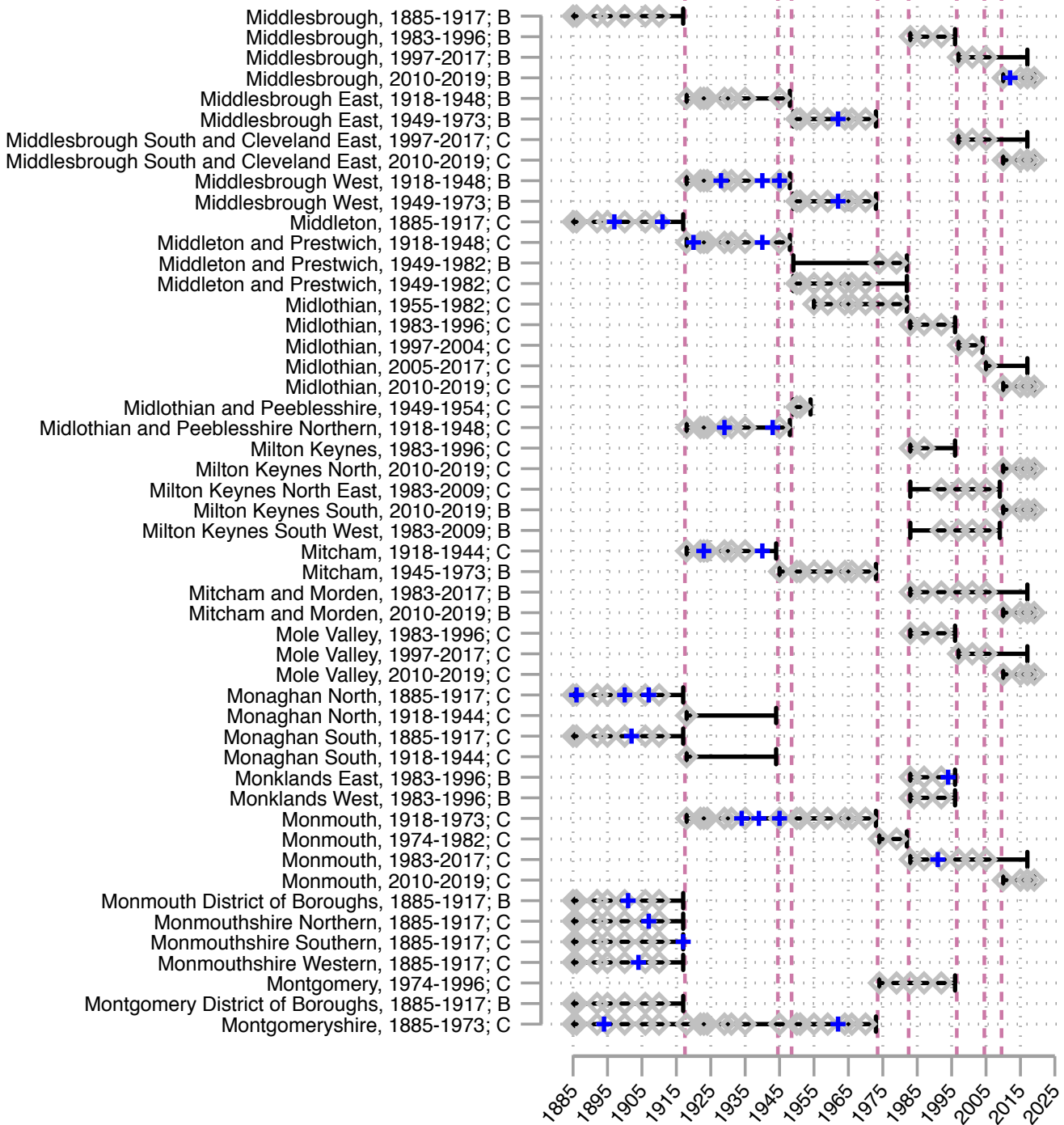
◇ General Election + By Election



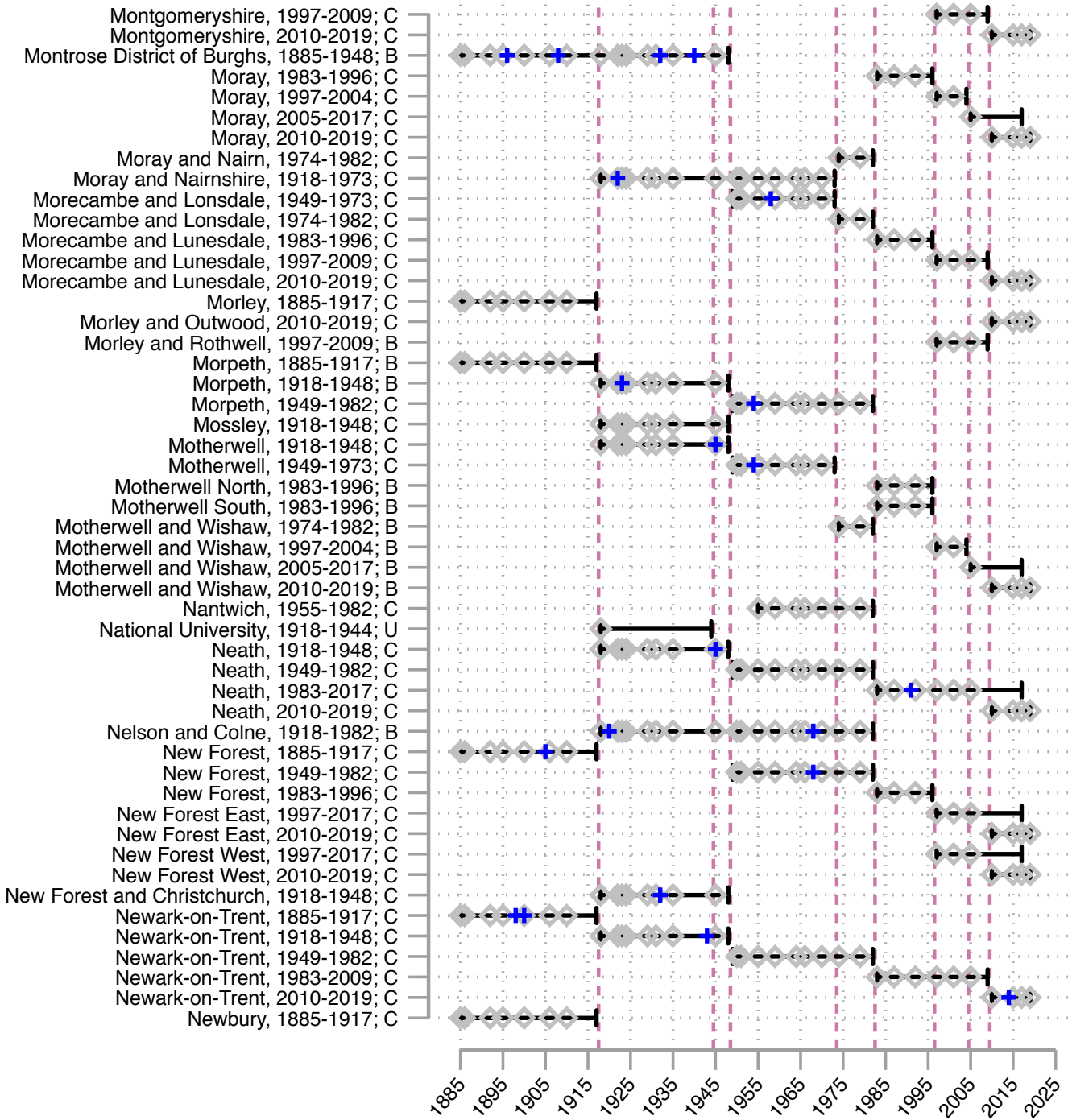
◇ General Election + By Election



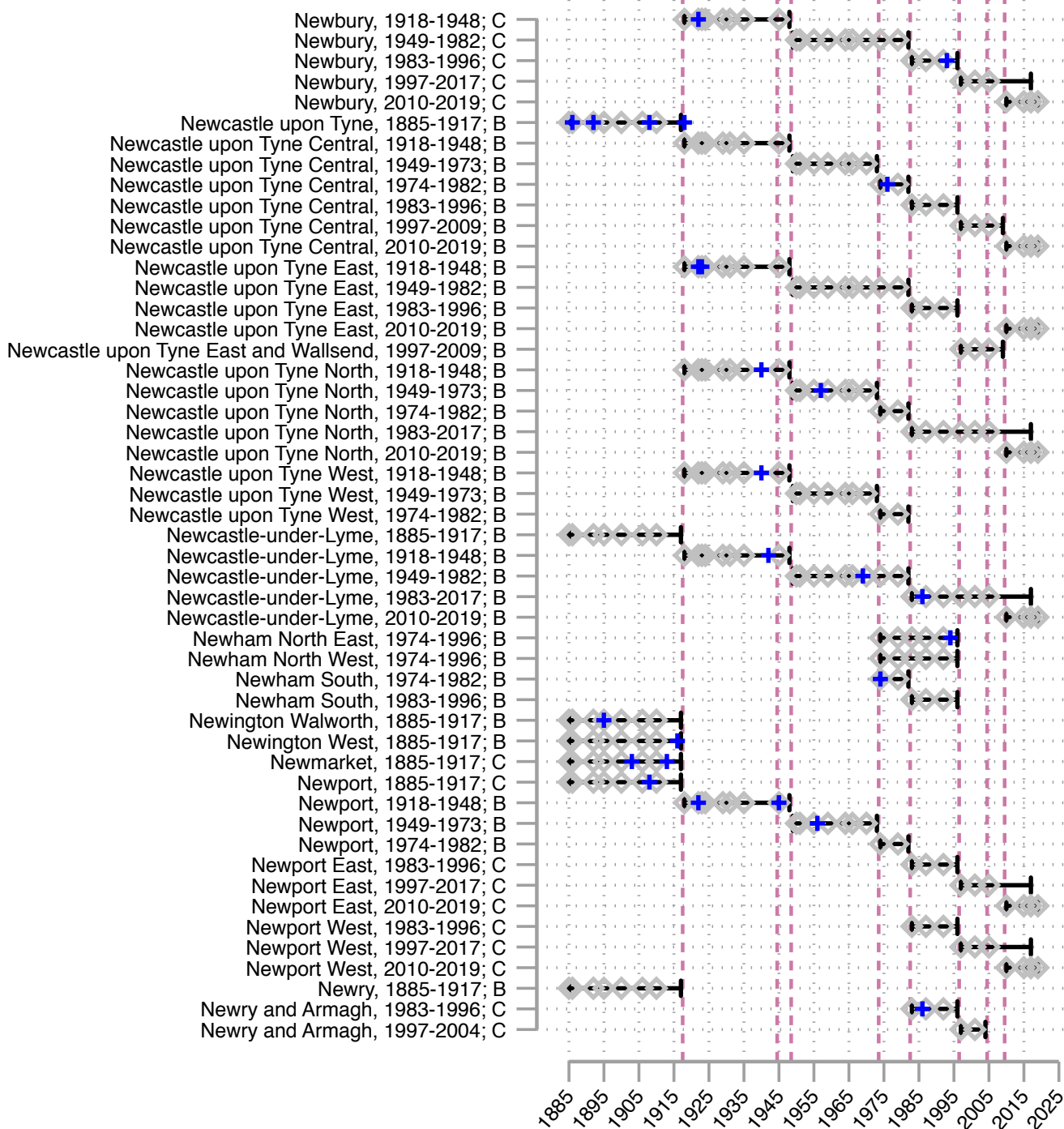
◇ General Election + By Election



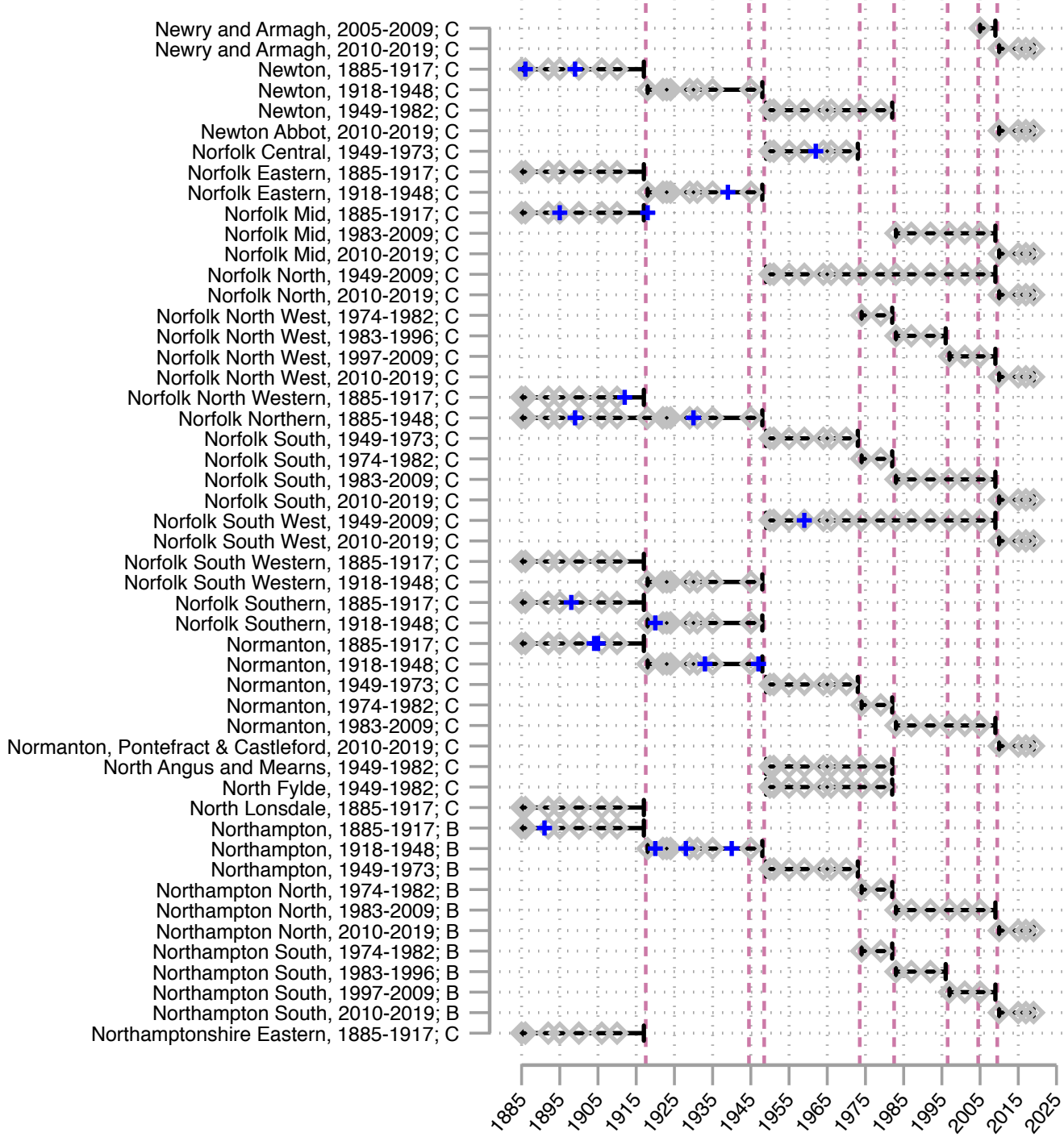
◇ General Election + By Election



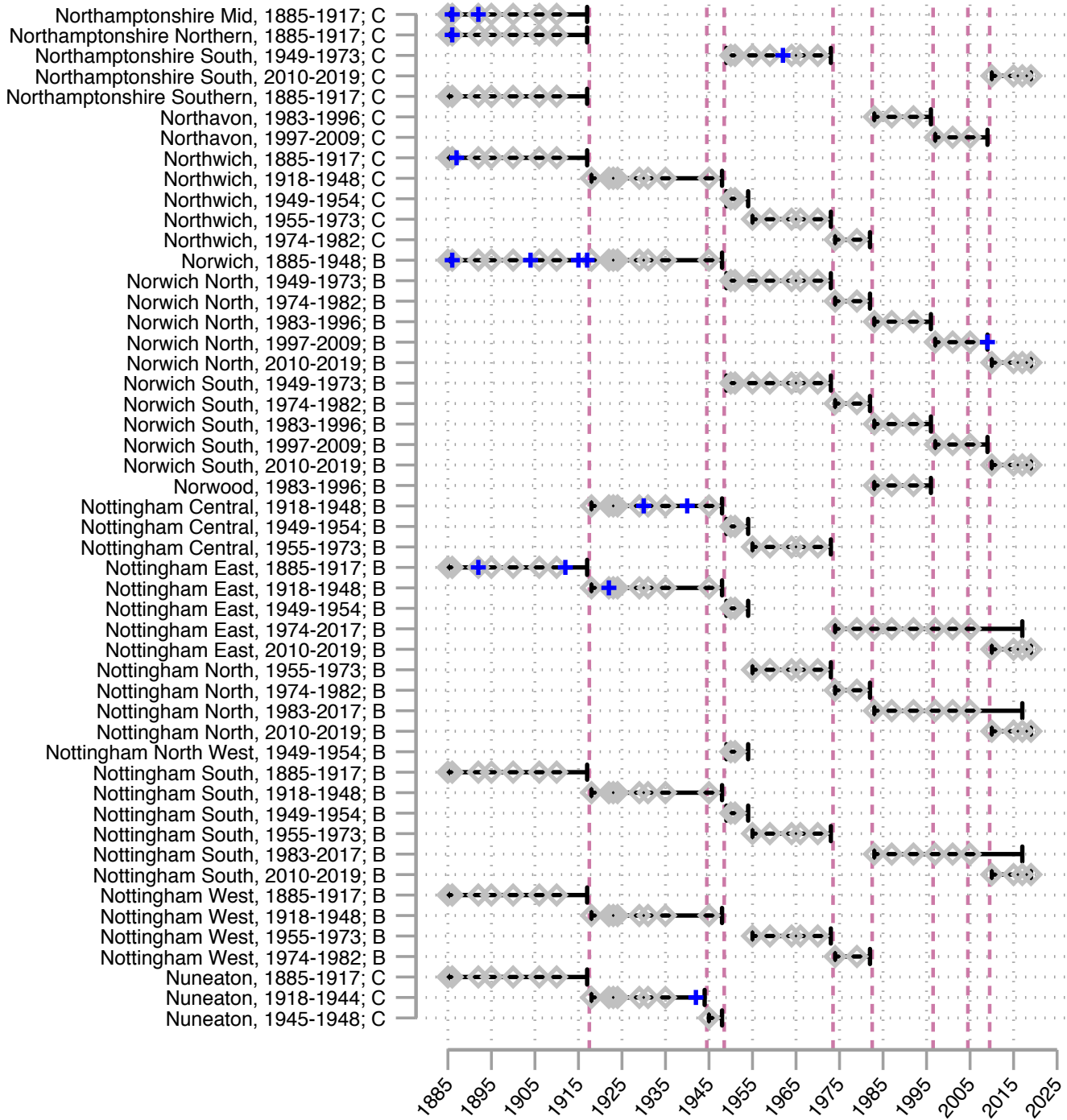
◇ General Election + By Election



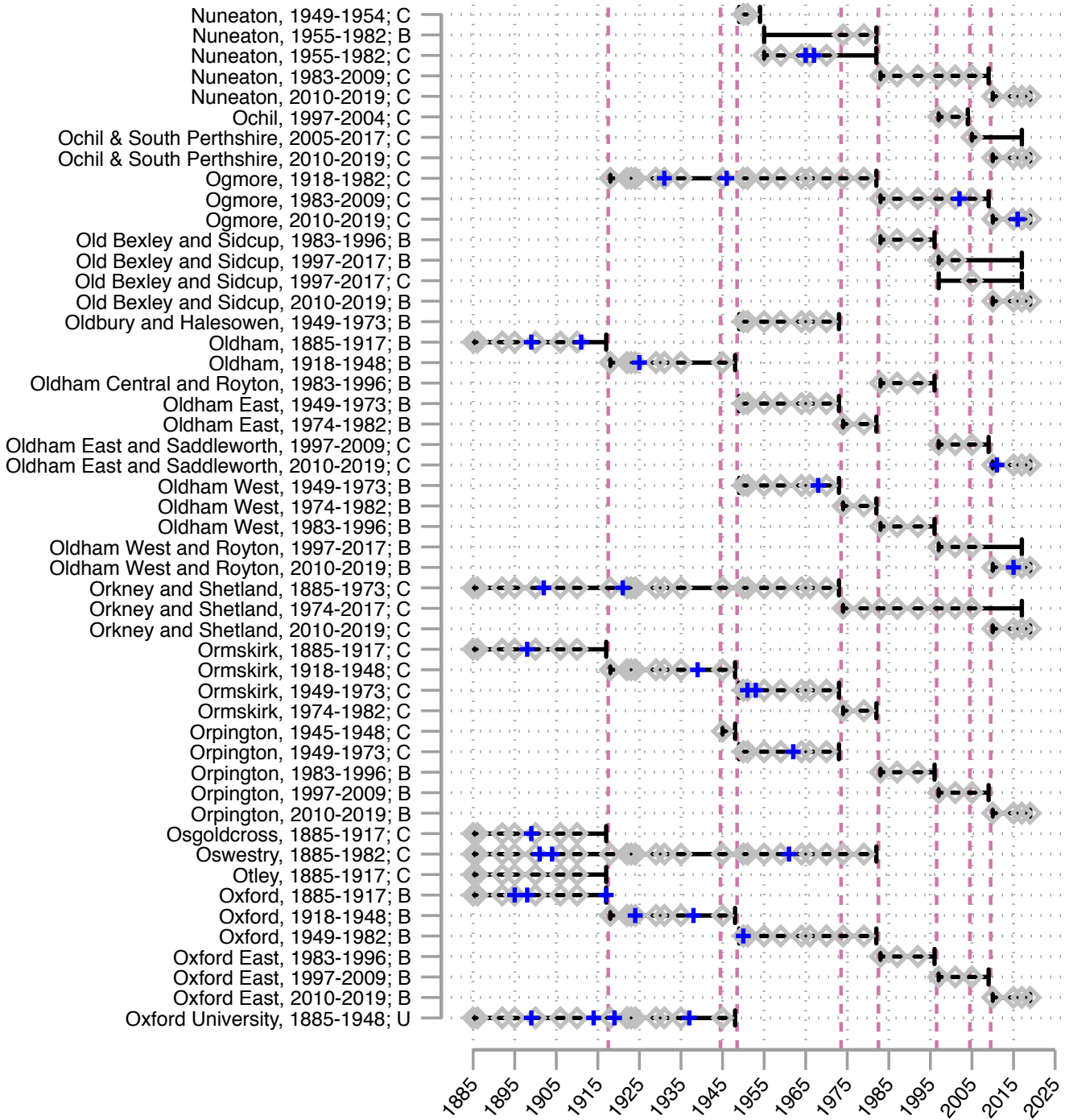
◇ General Election + By Election

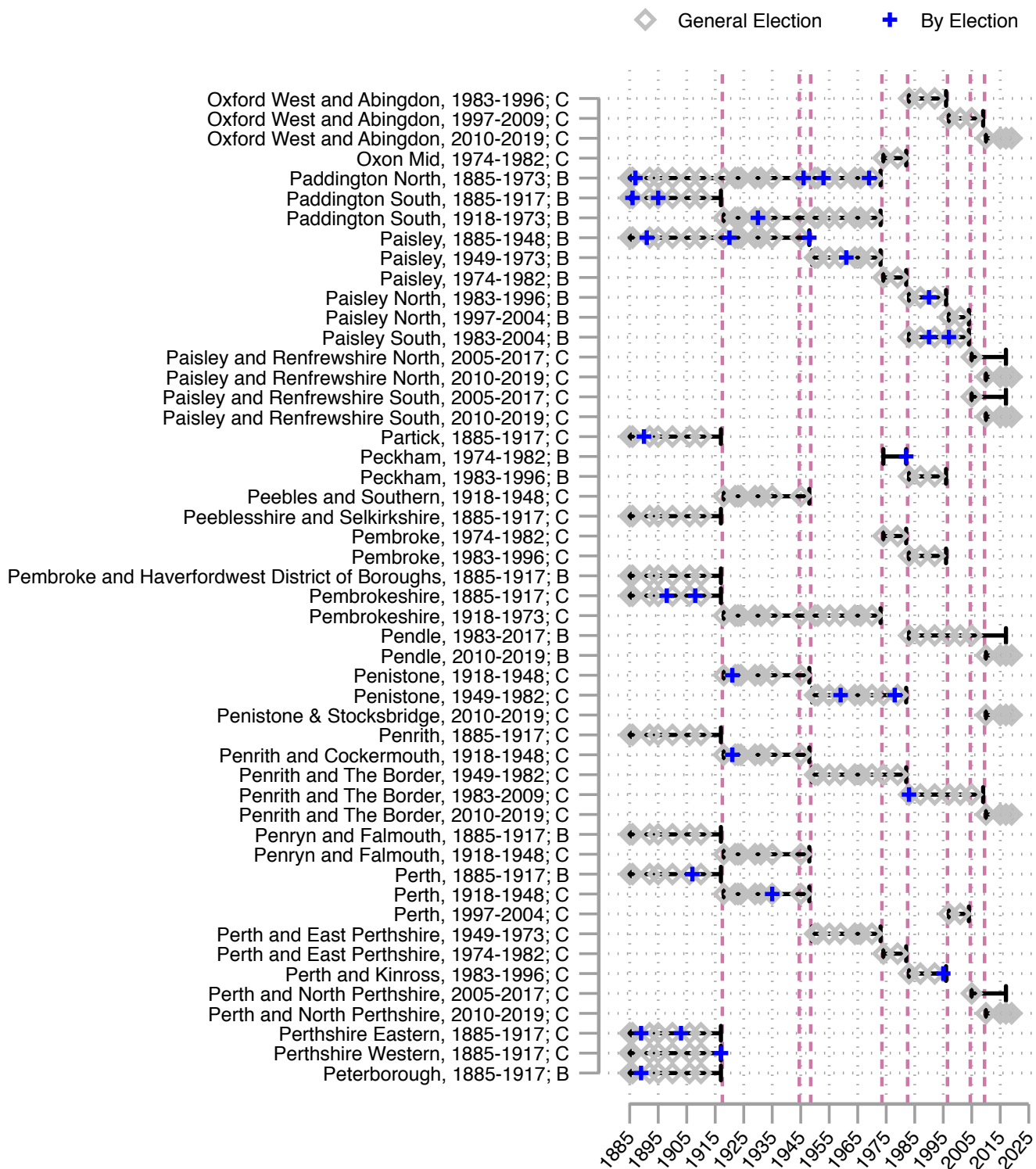


◇ General Election + By Election

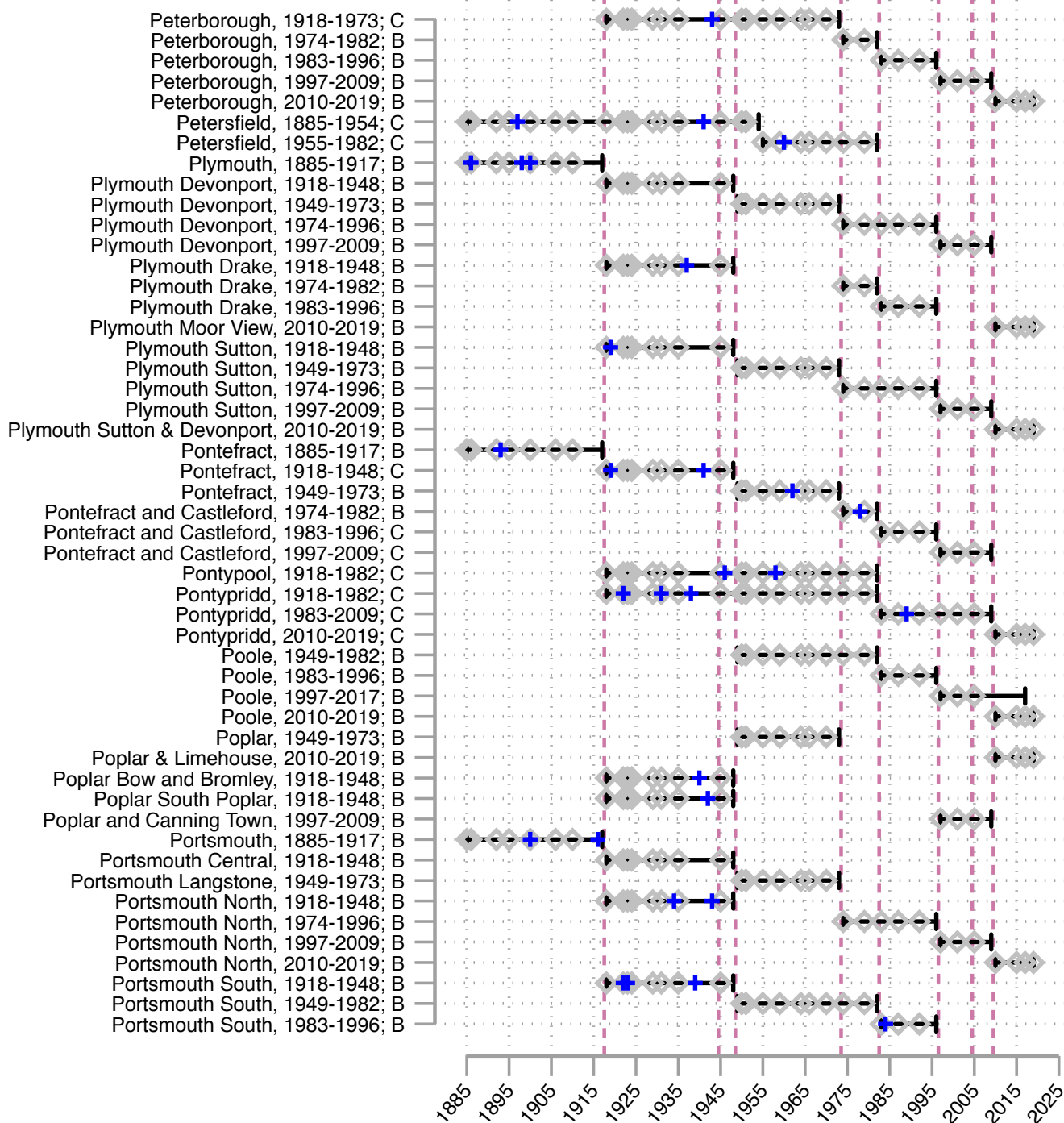


◇ General Election + By Election

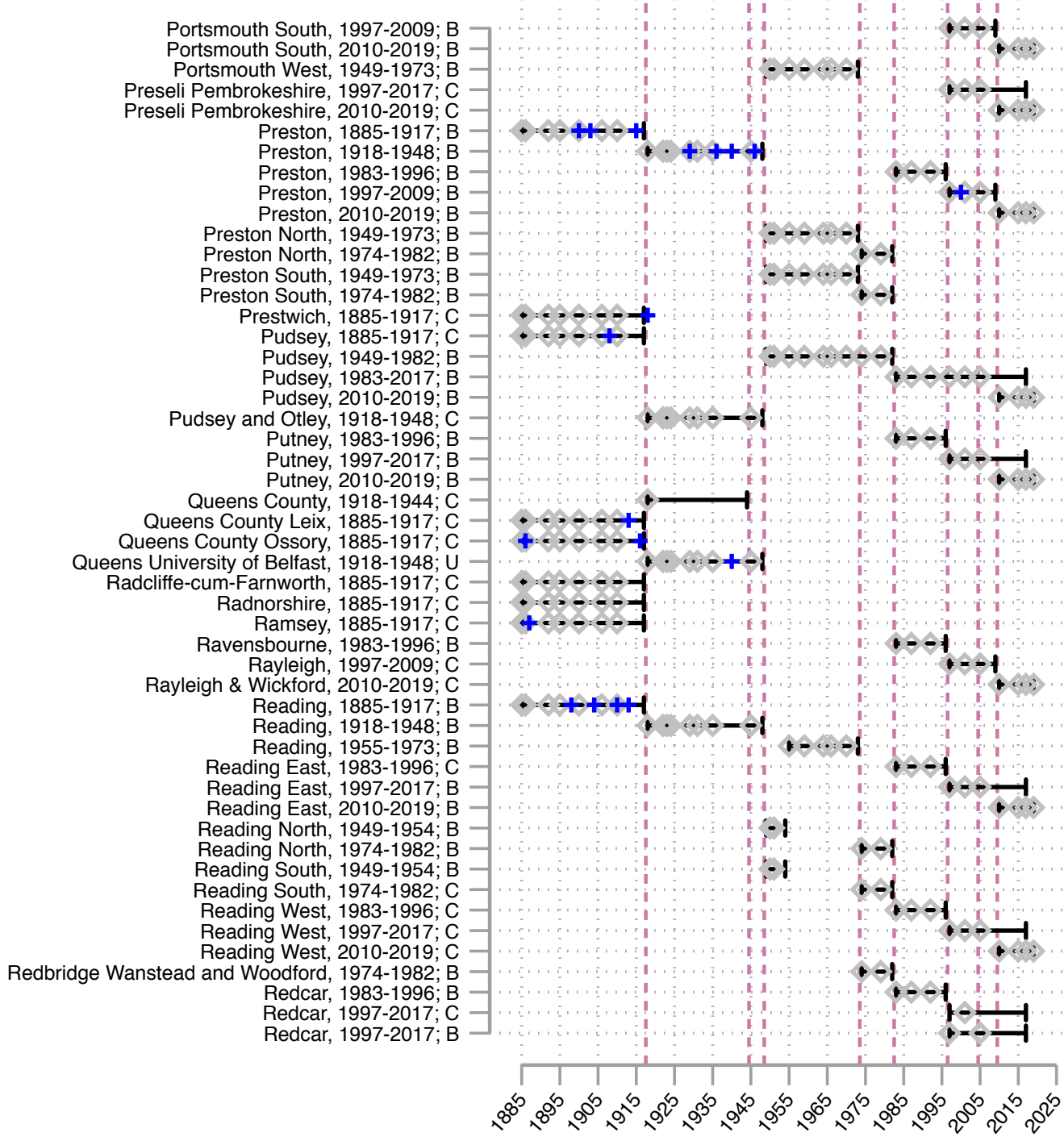


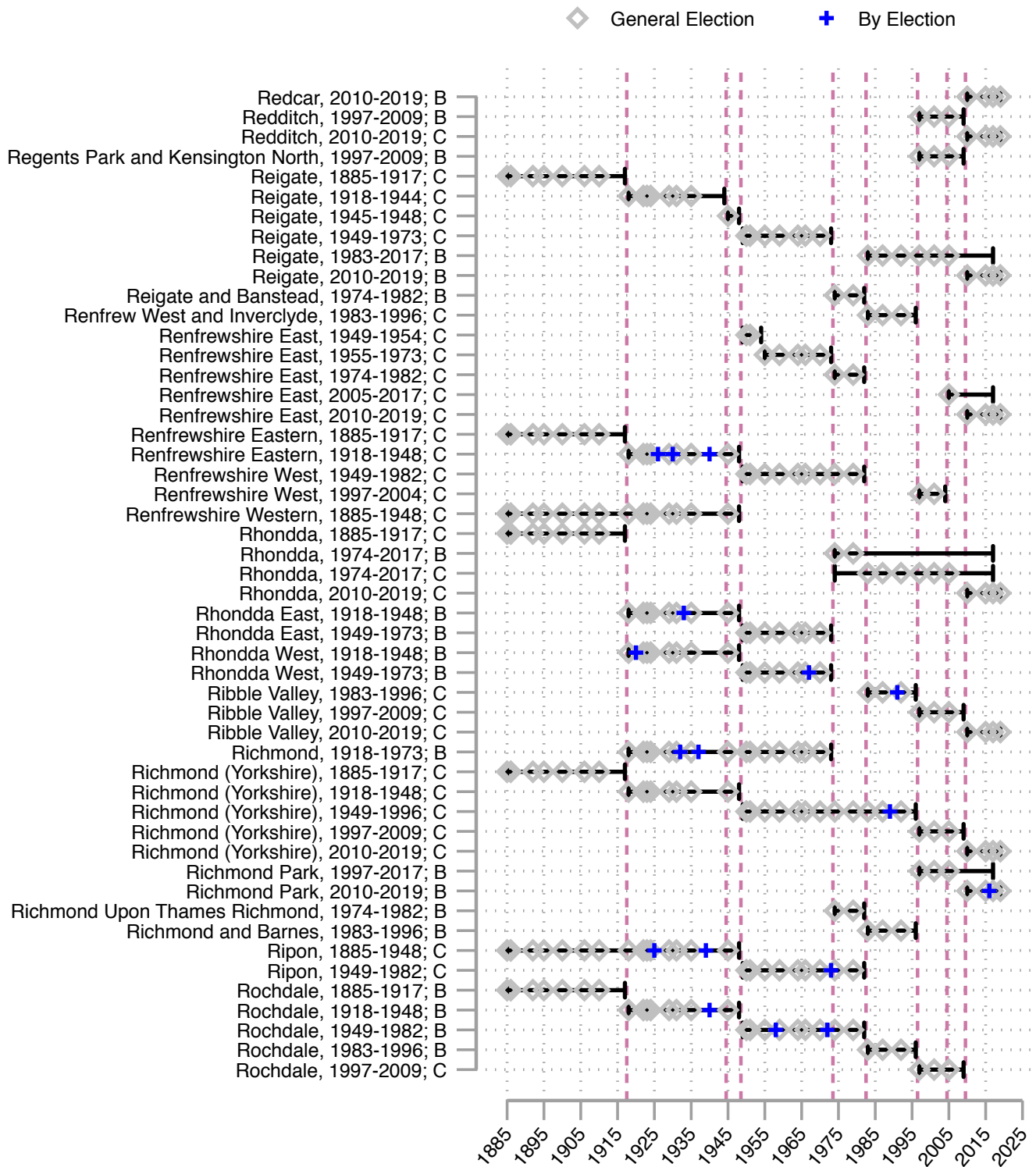


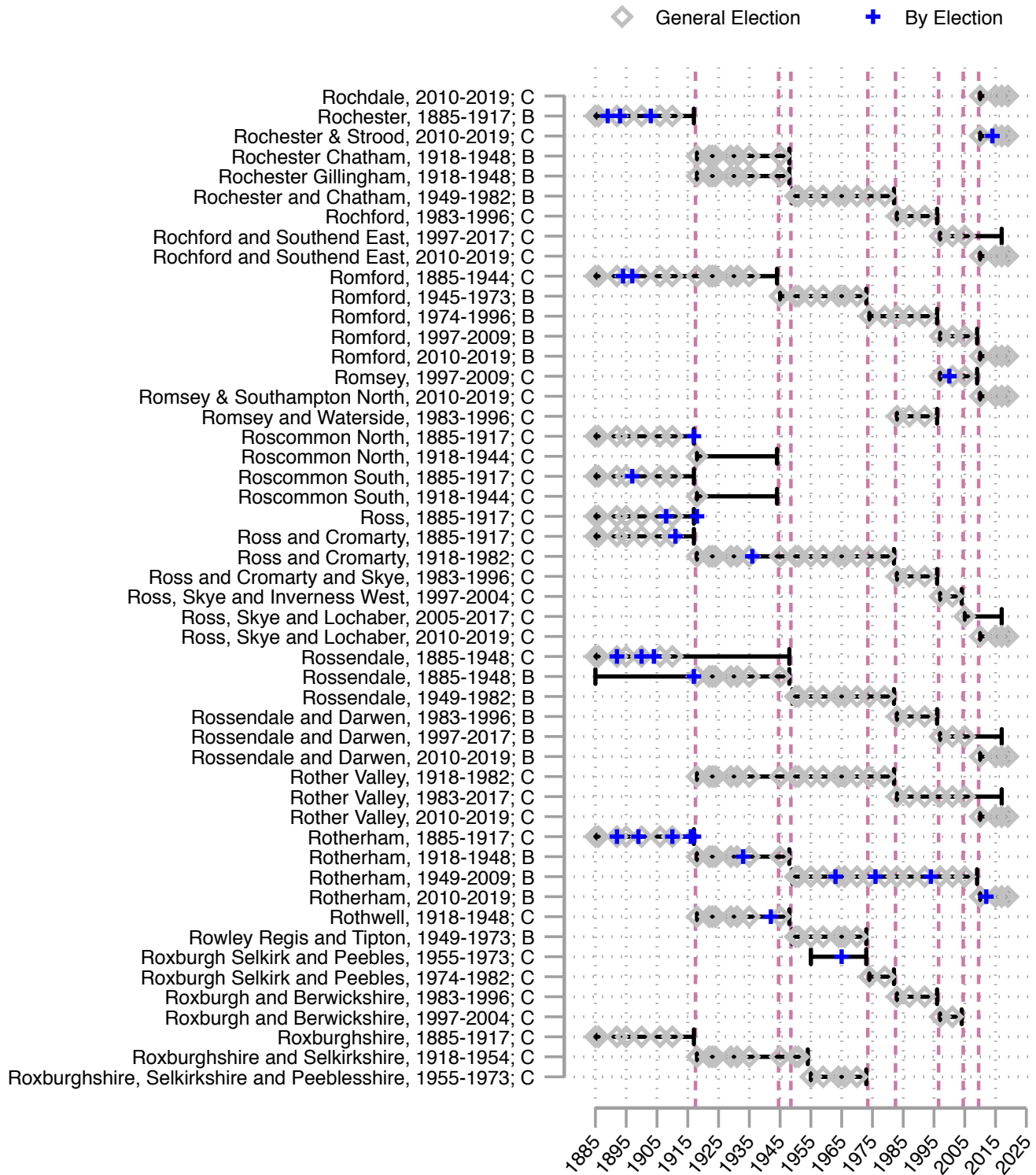
◇ General Election + By Election



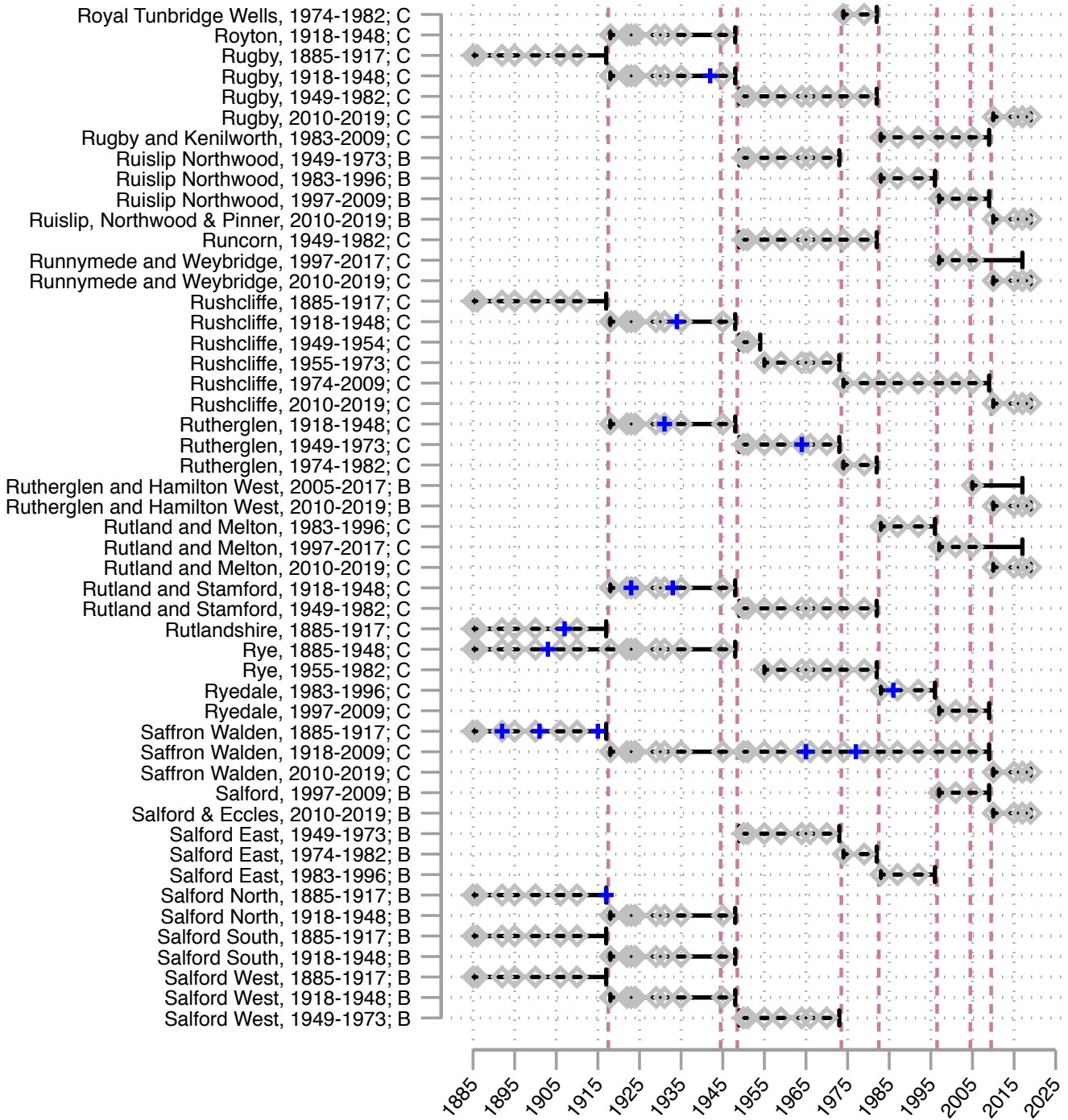
◇ General Election + By Election



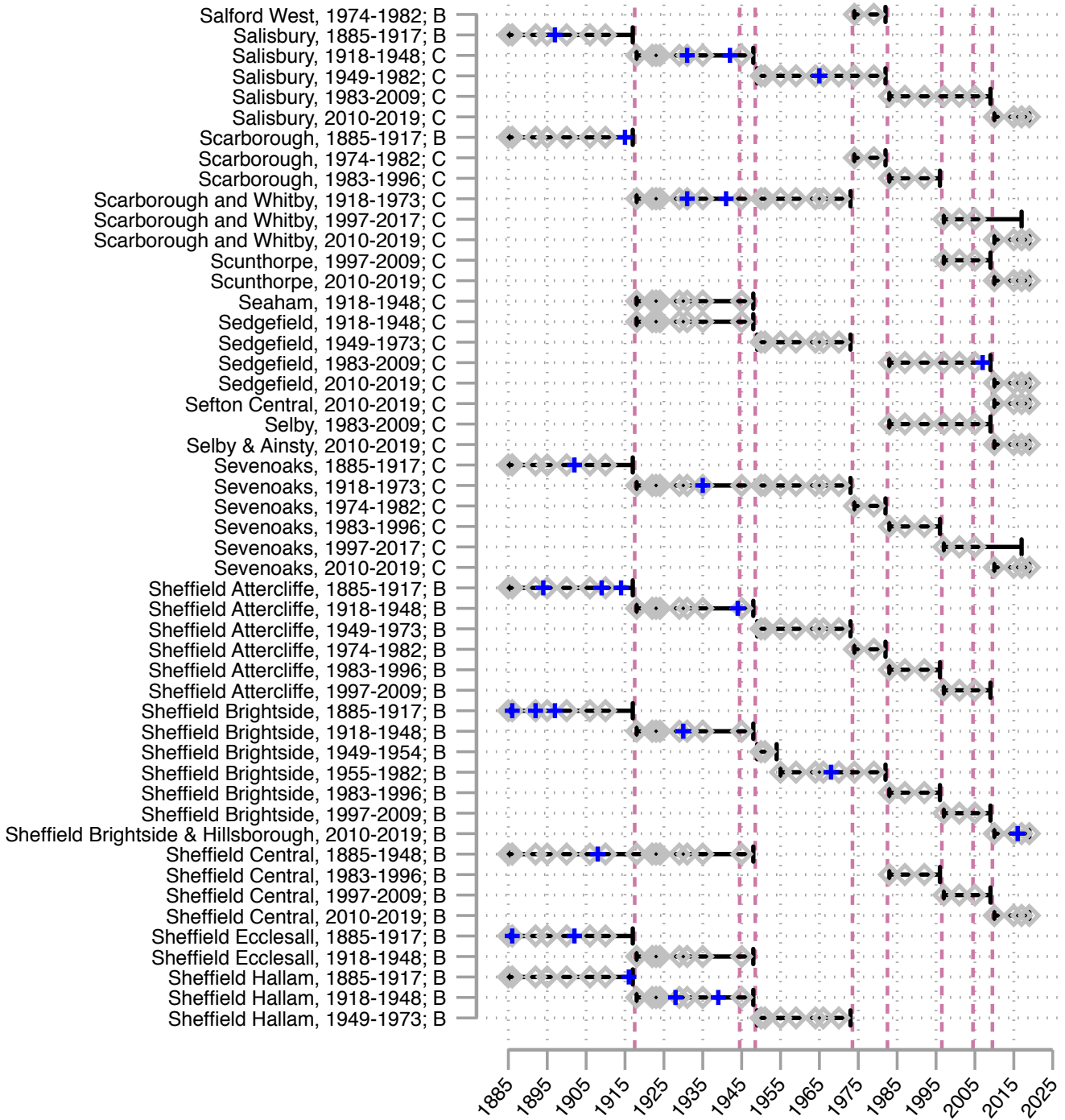




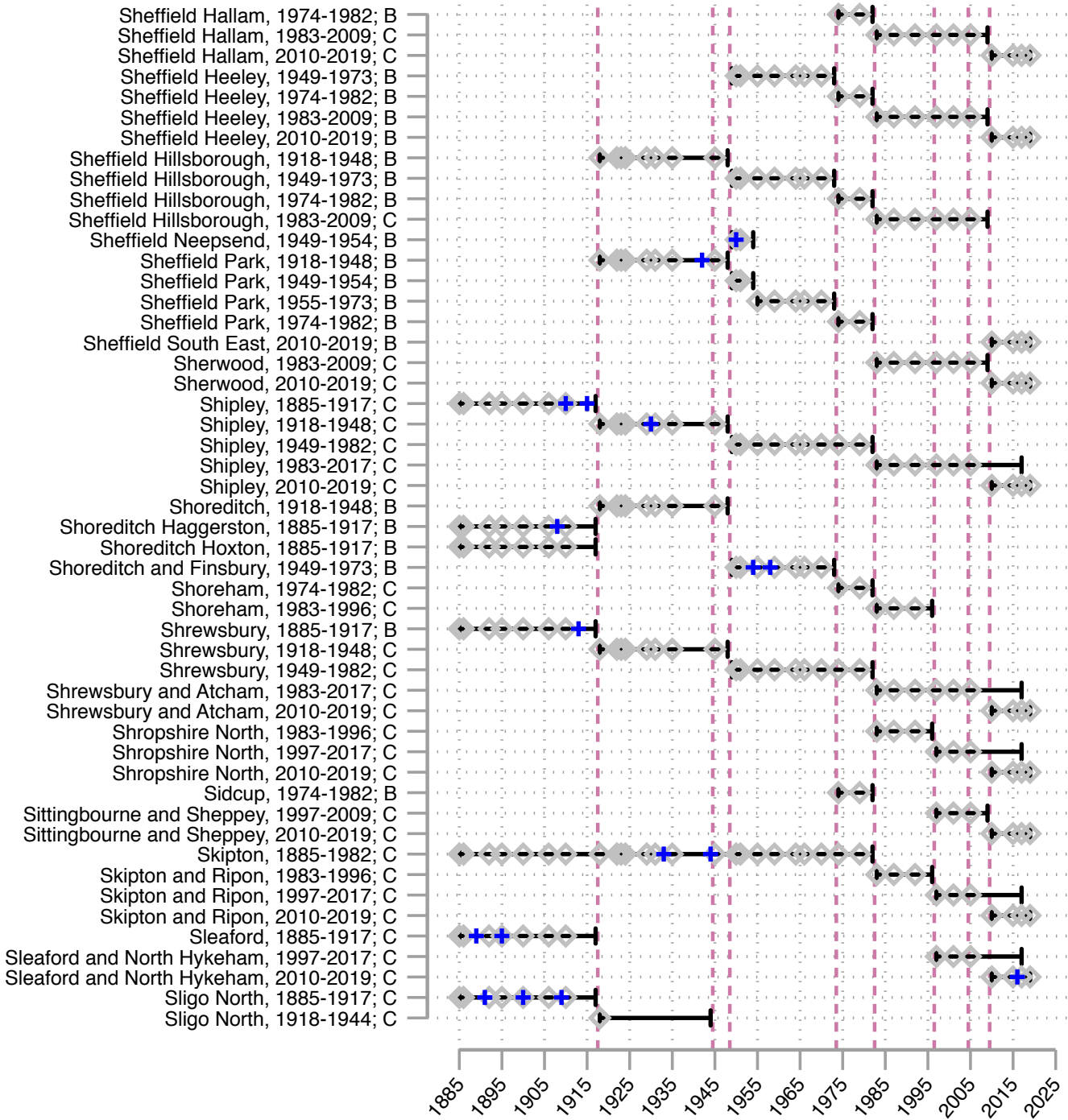
◇ General Election + By Election



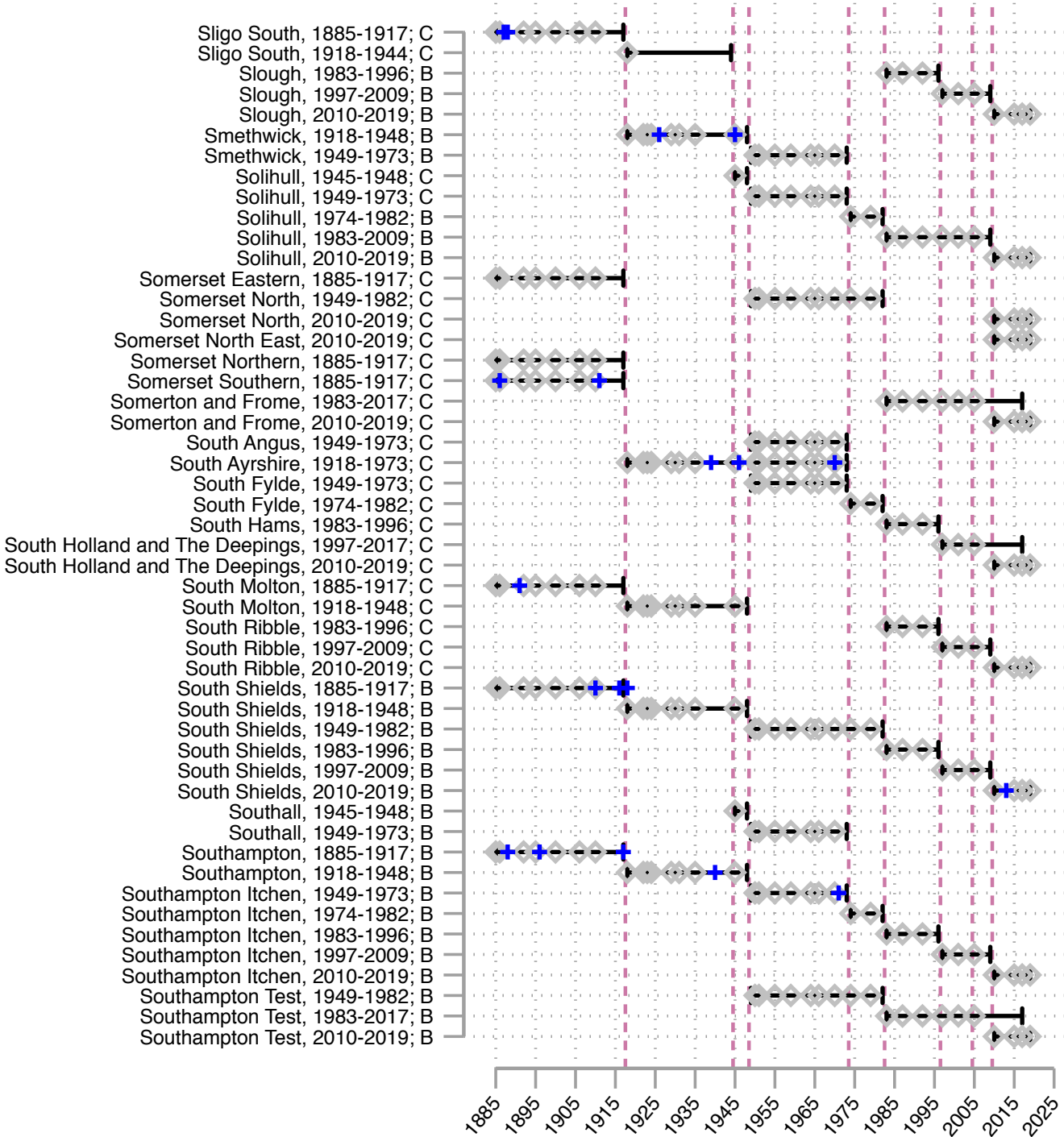
◇ General Election + By Election



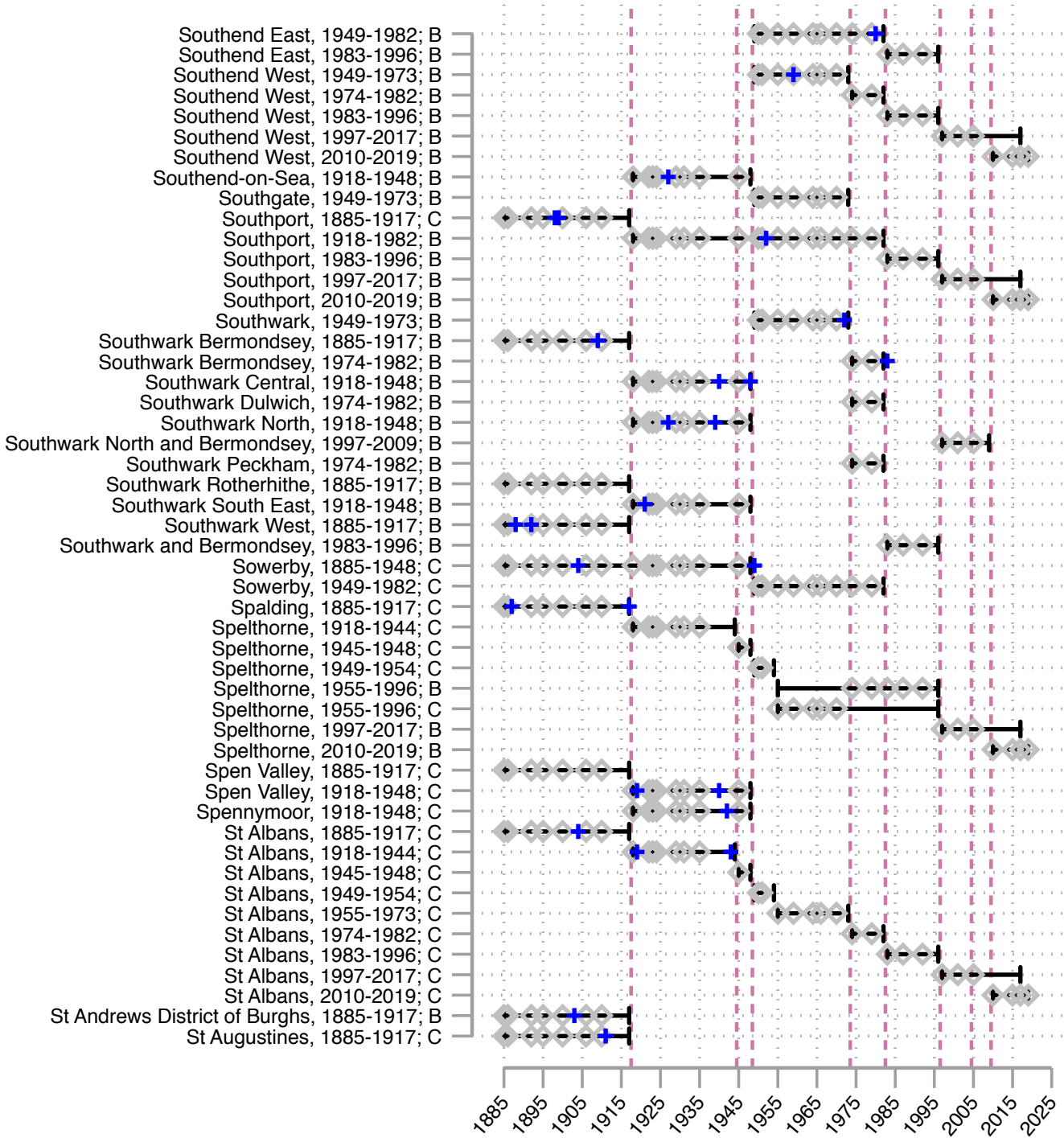
◇ General Election + By Election



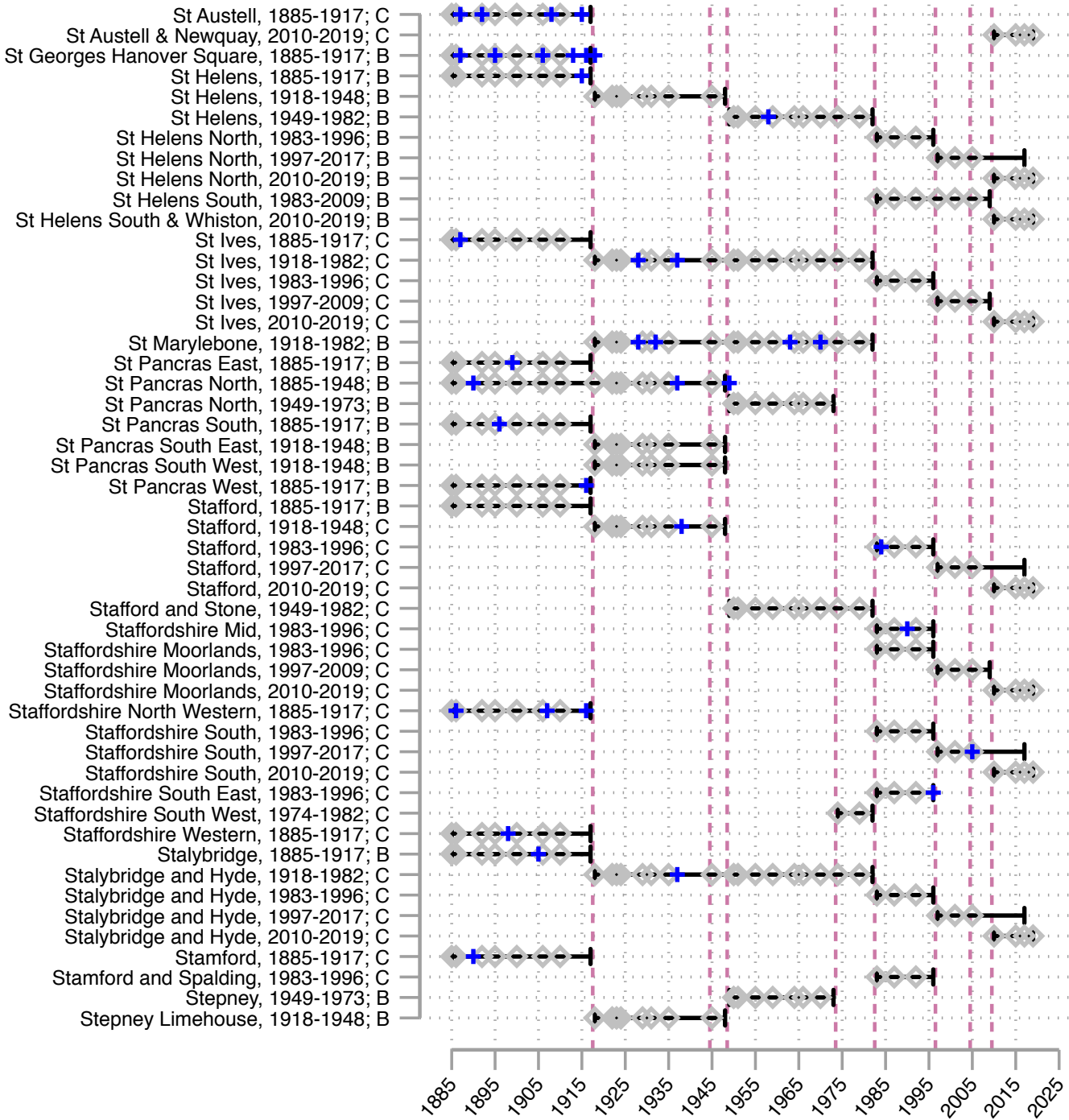
◇ General Election + By Election

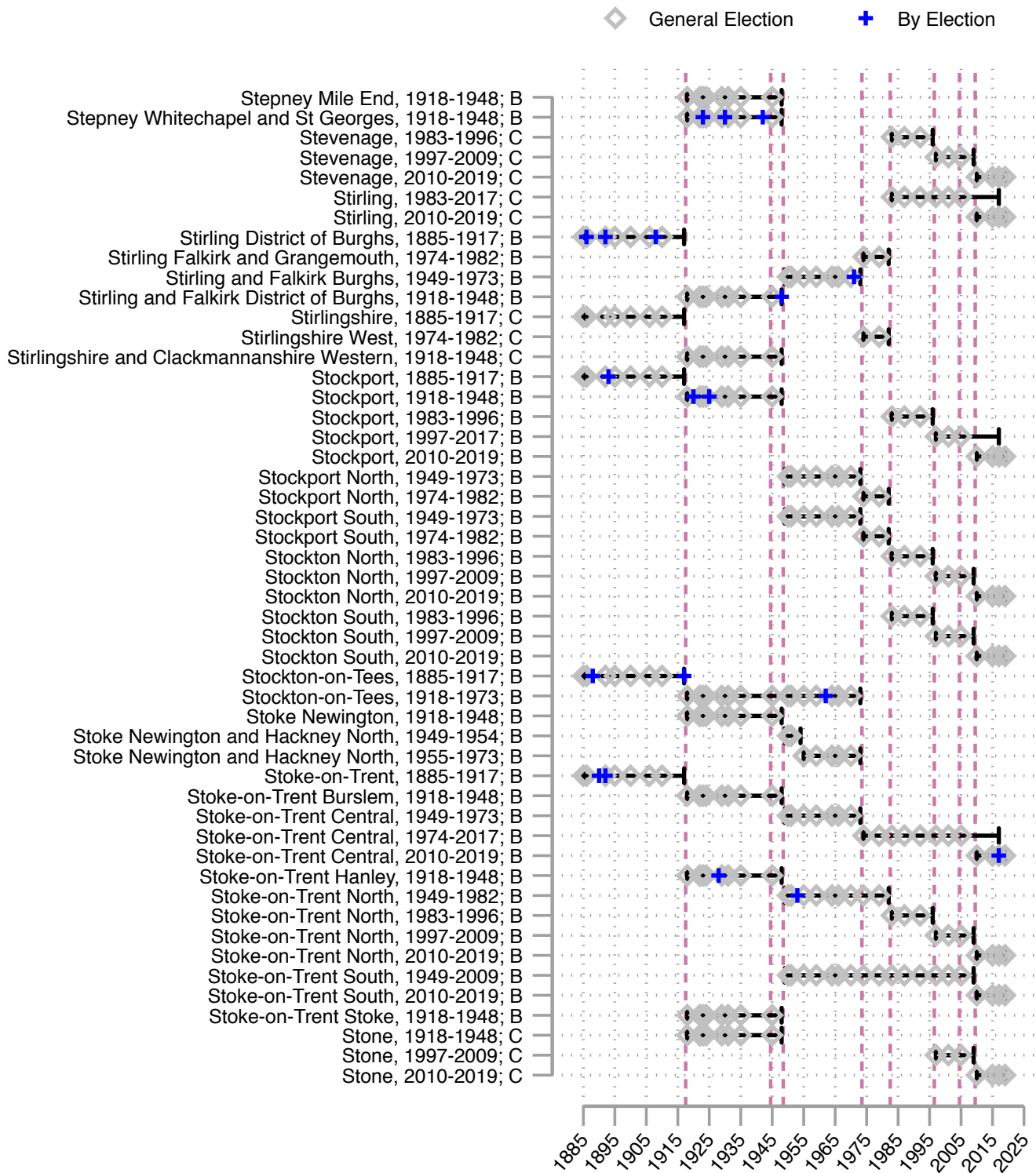


◇ General Election + By Election

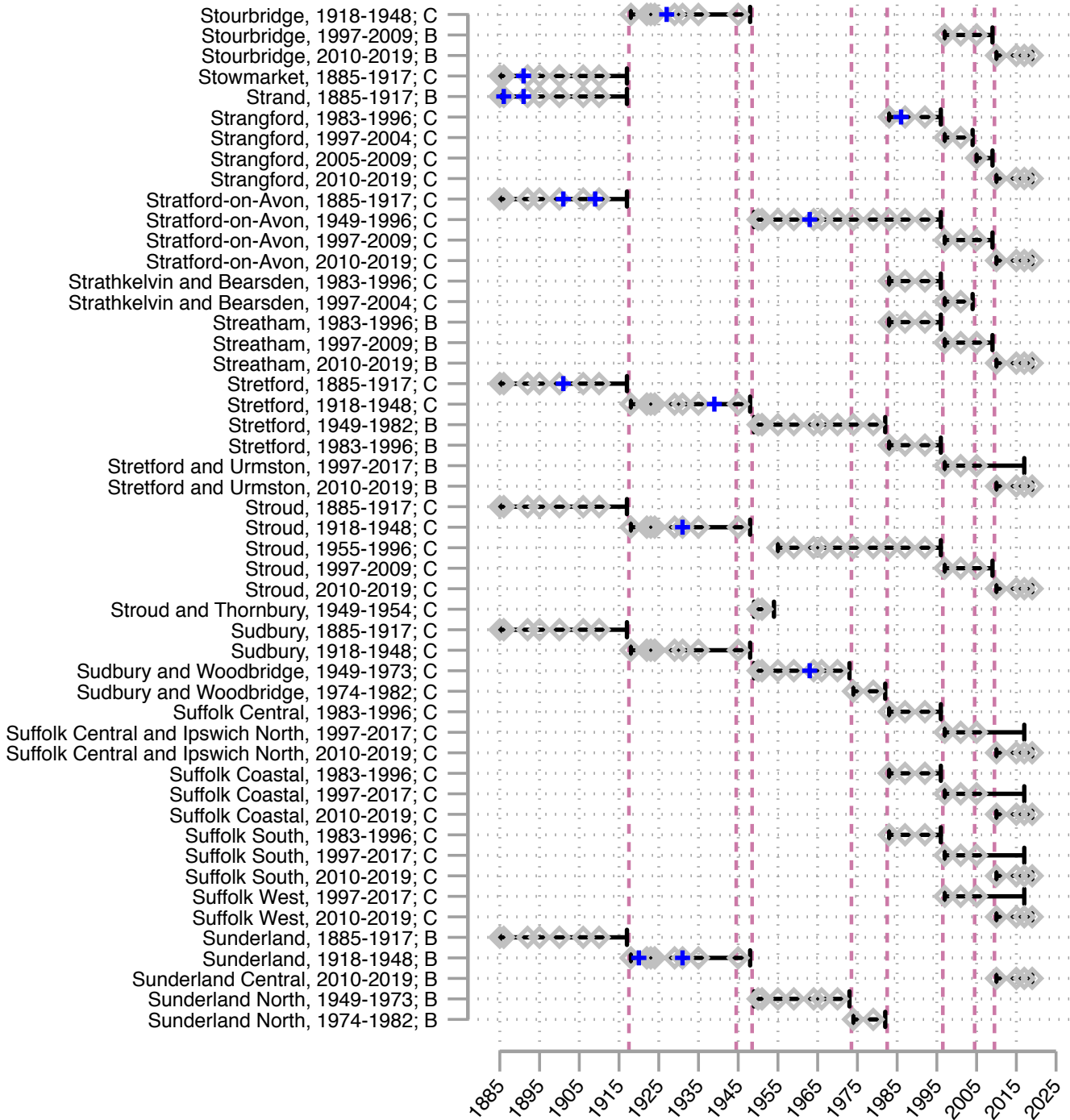


◇ General Election + By Election

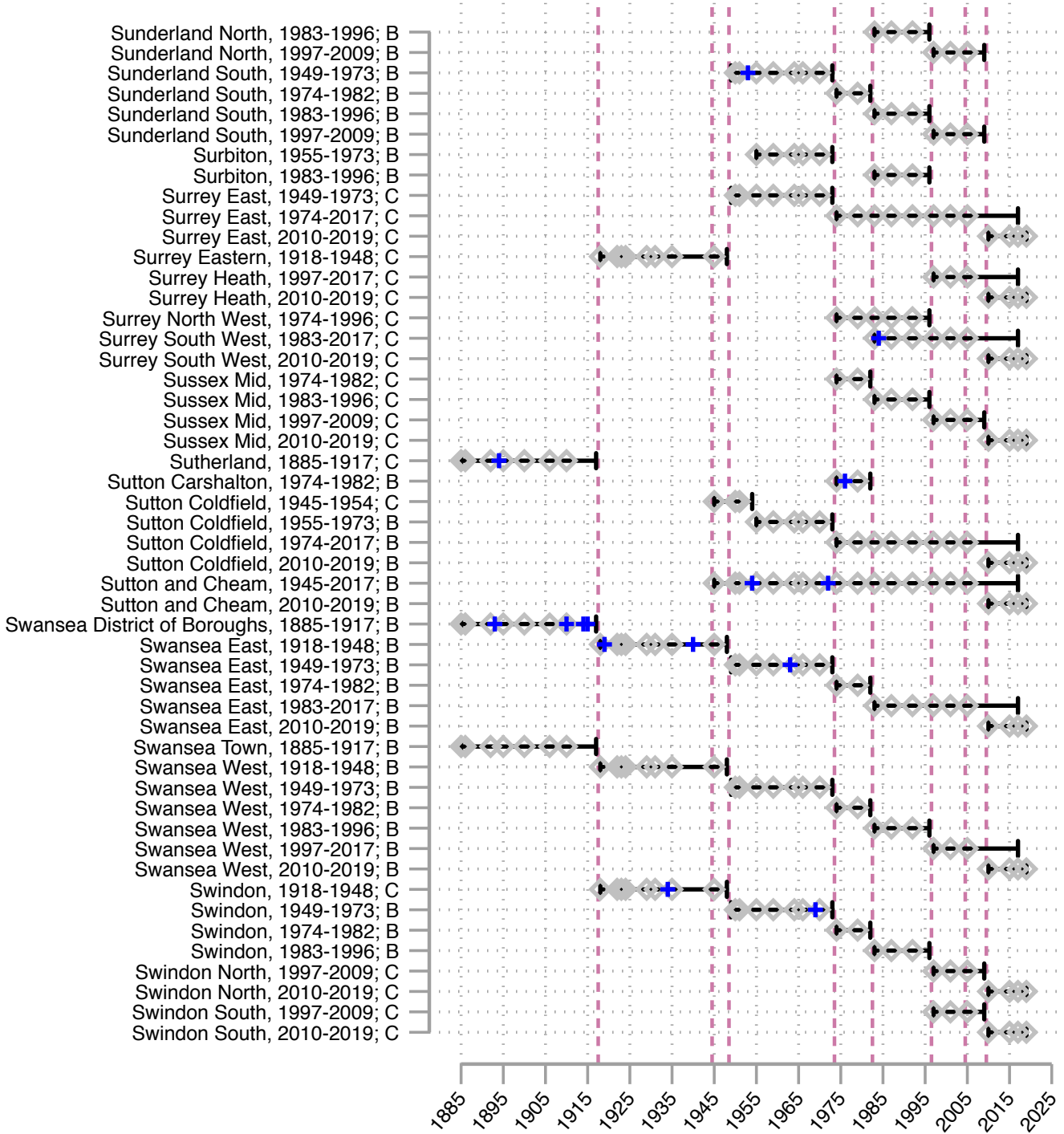




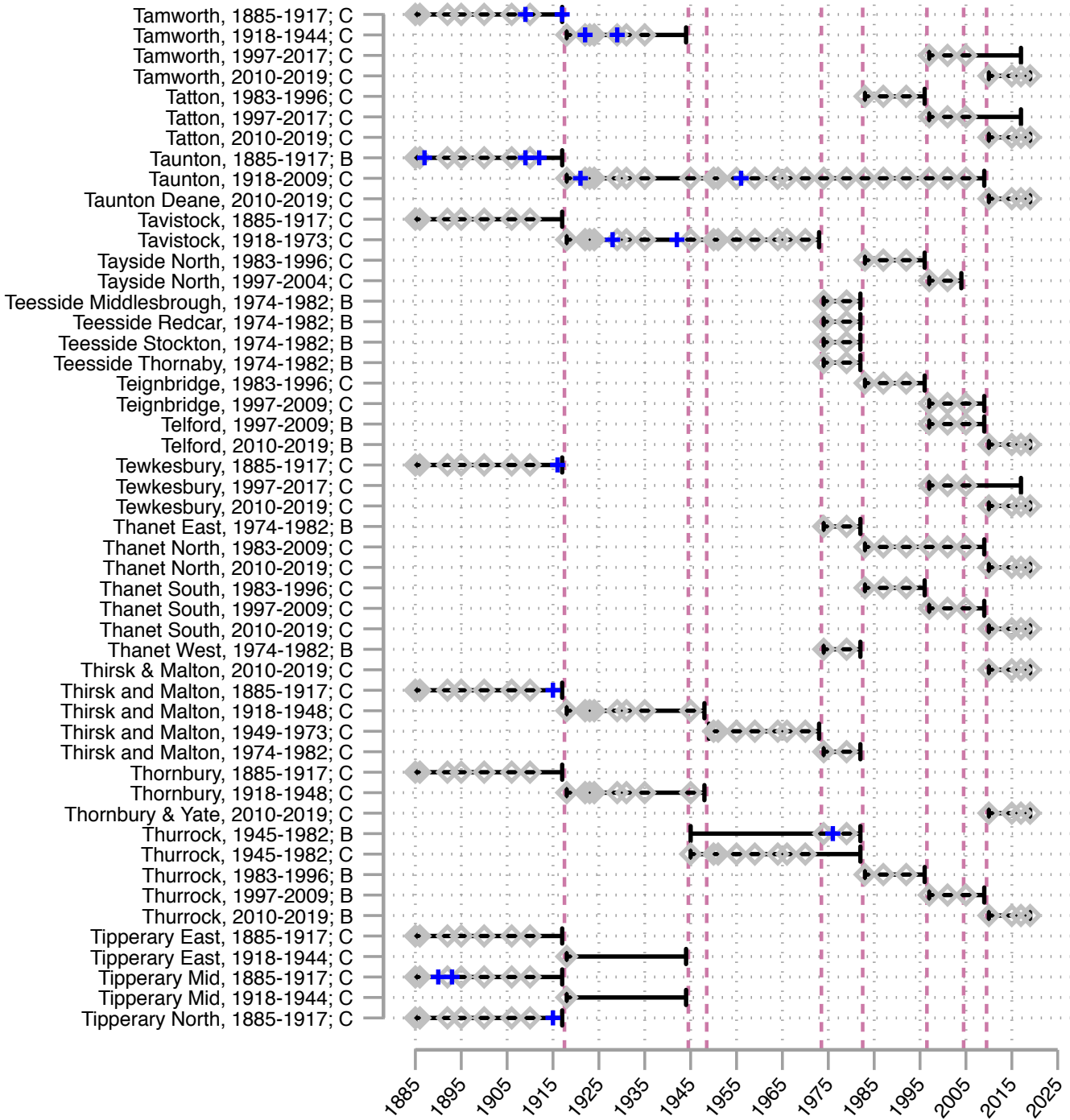
◇ General Election + By Election



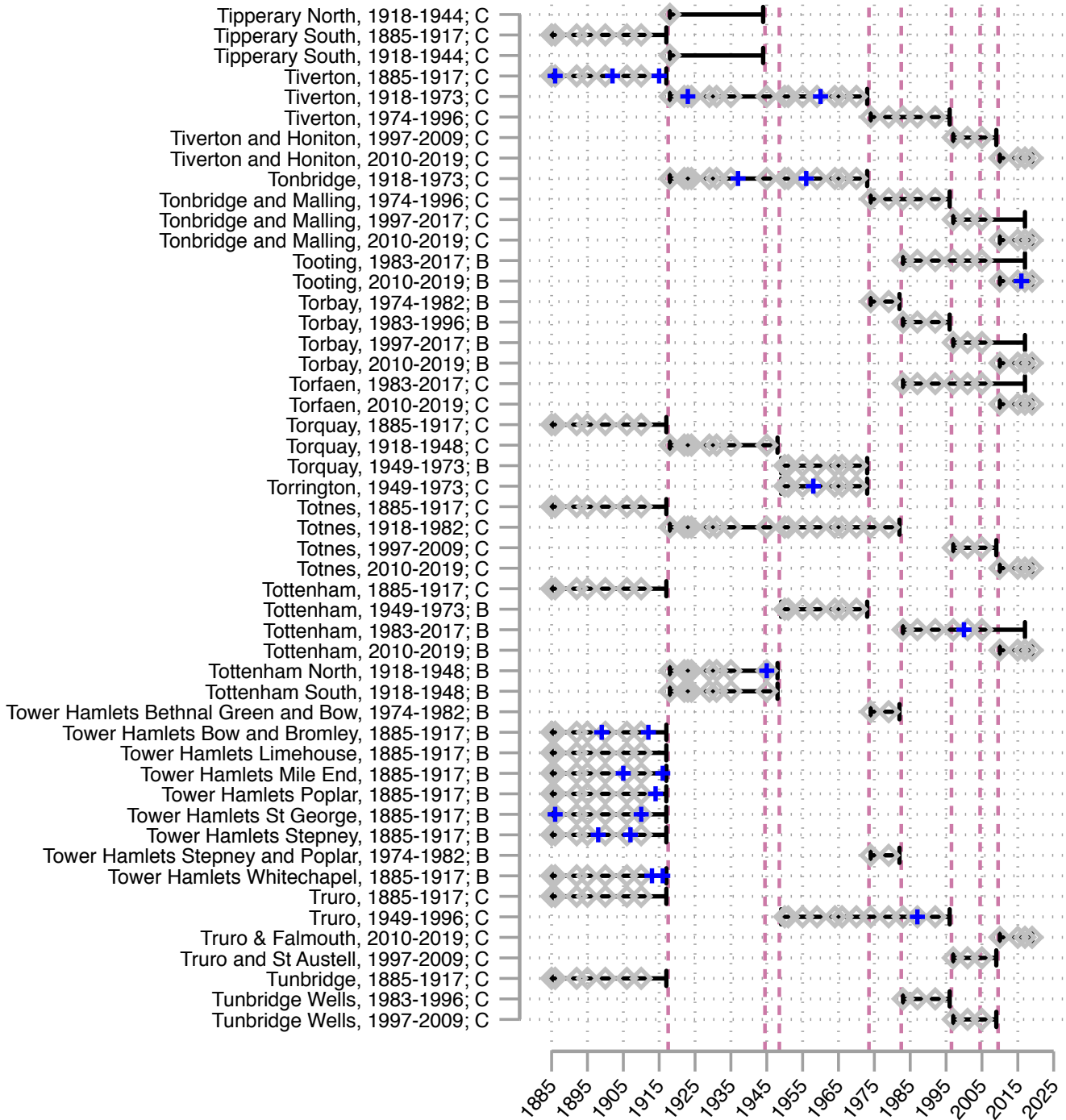
◇ General Election + By Election



◇ General Election + By Election



◇ General Election + By Election



◇ General Election + By Election

